Recombinant DNA

SECOND EDITION

James D. Watson
COLD SPRING HARBOR LABORATORY

Michael Gilman
COLD SPRING HARBOR LABORATORY

Jan Witkowski
BANBURY CENTER, COLD SPRING HARBOR LABORATORY

Mark Zoller
GENENTECH, INC.

SCIENTIFIC
AMERICAN
BOOKS

Distributed by

W. H. Freeman and Company

New York

The cover illustration, by Marvin Mattelson, symbolizes some of the elements of this book. The DNA double helix is, of course, central to the book, as it is to the cover illustration. The blocks are double-stranded DNA fragments synthesized by the polymerase chain reaction, a technique that has revolutionized the way molecular genetics experiments are done. The number of fragments doubles repeatedly, going off into the distance (see Chapter 6). The coat colors of the mice running down the helix (in the same direction but with opposite polarity!), are changing from albino to chimeric, then chimeric to agouti. These coat color changes show mice in which genetic engineering has been used to knock out a specific gene. The experiment is shown more realistically in Figure 14-9.

Library of Congress Cataloging-in-Publication Data.

Recombinant DNA/James D. Watson... [et al.]. — 2nd ed.
p. cm.
Includes bibliographical references and index.
ISBN 0-7167-1994-0. — ISBN 0-7167-2282-8 (pbk.)
1. Recombinant DNA. I. Watson, James D., 1928–
QH442.R37 1992
574.87'3282—dc20

91-38483
CIP

Printed in the United States of America

Scientific American Books is a subsidiary of Scientific American, Inc.

Distributed by W. H. Freeman and Company, 41 Madison Avenue, New York, New York 10010

1 2 3 4 5 6 7 8 9 0 RRD 9 9 8 7 6 5 4 3 2

Contents

Analysis of Cloned Genes

New Tools for Studying Gene Function

Impact of Recombinant DNA on Human Genetics

Preface

Application of recombinant DNA techniques to biology is bringing about a revolution in our understanding of living organisms. There is no field of experimental biology that is untouched by the power we now have to isolate, analyze, and manipulate genes. When the first edition of *Recombinant DNA* was published in 1983, recombinant DNA techniques were already being used extensively for the analysis of viral and bacterial genetics, but dissection of eukaryotic genes was only just beginning. There were hints of what was to come. The concept of the gene as a continuous stretch of DNA had been shattered with the discovery of introns, but alternative splicing and genes-within-genes were yet to be revealed. Identification of cellular oncogenes seemed to promise an understanding of cancer, but the mechanisms of their action—and the existence of tumor suppressor genes—were still subjects for speculation. A handful of genetic diseases were being analyzed at the molecular level, but the isolation of the disease genes and the development of gene therapy were yet to come.

Our aim in writing the second edition of *Recombinant DNA* is to show how recombinant DNA techniques have led to the explosion in our knowledge of fundamental biological processes. As in the first edition, which was subtitled *A Short Course*, we provide a concise presentation of the methods, underlying concepts, and far-reaching applications of recombinant DNA technology. The field has grown since the publication of the first edition, and so has our book. But even though our previous subtitle may be inappropriate for this enlarged edition, our approach to the material has remained true to the spirit of the "short course": as before, the uninitiated will find access to the field of recombinant DNA here.

The book is now divided into six major sections. The first five chapters, which are largely unchanged from the first edition, provide a historical introduction to the early development of recombinant DNA technology, up to the point when studies of eukaryotic organisms began in earnest. In the next section we describe in detail the methods currently used to clone and analyze genes, and devote an entire chapter to the polymerase chain reaction, which has had an extraordinary impact on research. The great power of recombinant DNA techniques comes from the ability to explore gene functions by manipulating genes and then introducing them back into cells. The third section of the book discusses how this is done in mammalian cells, yeast, mice, and plants. The fourth section describes the progress these manipulations have allowed in key areas of biology. Here the range of recombinant DNA applications is demonstrated, from the analysis of cell cycle control and embryonic development, to the isolation of genes involved with brain function. Indeed, these techniques have spawned a whole industry—biotechnology. In the fifth section, we describe some of its accomplishments, including the development of genetically engineered pharmaceutical and agricultural products, and the studies of the human immunodeficiency virus that are leading the attack on AIDS. The differences between the first and second editions are perhaps most evident in the final section, where we describe the revolution in human molecular genetics and the ways in which recombinant DNA techniques are providing new methods for diagnosis and treatment of human inherited diseases.

The topics that are covered and the approach we take to describing them make this book suitable for undergraduate and graduate students in molecular biology, cell biology, biochemistry, genetics, or biotechnology courses; for medical students and physicians; and for others who have an interest in recombinant DNA techniques—for example, forensic scientists, patent attorneys, and science journalists.

Textbooks dealing with biochemistry, molecular genetics, and molecular biology usually present information without describing the experiments that were done to obtain it. We think that this is a pity, because designing and doing experiments is exciting and fun. As in the first edition, we have used real experiments to illustrate important biological phenomena, and we have plundered our colleagues' papers for interesting examples. Figures are used profusely to try to make complex real-life experiments intelligible, but inevitably we have not been able to present all the subtle details. Those who want to explore these details will find the experiments in the research papers listed at the end of each chapter, and the review papers we cite will provide an entry point to each topic.

This book is atypical in another regard. Because we do not consider it primarily a textbook for conveying undisputed facts about molecular biology, we have been able to include exciting research that is at the cutting edge of biology. The interpretation of experimental data often changes with time, so the reader should bear in mind that future research might require modification of some of the ideas we present. This is all part and parcel of doing research, because a science that does not change is a dead science. Modern experimental research in biology is an ever-changing dynamic enterprise, and we hope that *Recombinant DNA* conveys the excitement of the continuing process of discovering how organisms work.

Acknowledgments

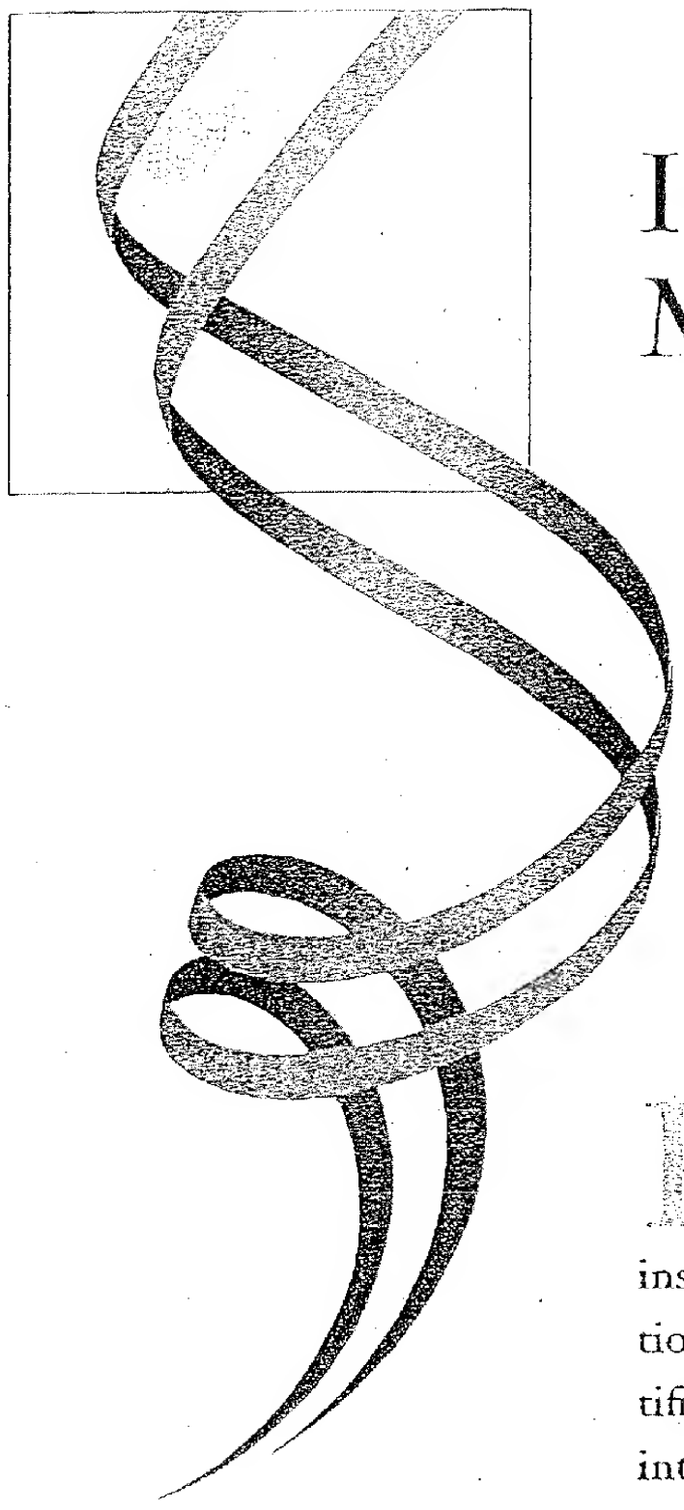
Those who most deserve our thanks are our families, from whom we were taken for many days by this book, and our colleagues at Cold Spring Harbor Laboratory and Genentech, who sometimes had a difficult time communicating with us when we were preoccupied with writing.

We are grateful to the many friends and colleagues who read our manuscript, criticized, corrected, and provided information. They include Sue Alpert, French Anderson, Avi Ashkenazi, David Beach, Martin Bobrow, Tom Caskey, Jeff Chamberlain, Irvin Chen, Francis Collins, Alan Coulson, David Cox, Ken Culver, Kay Davies, Jim Eberwine, Stan Fields, Ted Friedman, Bruce Futcher, Peter Gergen, Richard Gibbs, Paul Godowski, Andre Goffeau, Takashi Gojobori, Kenshi Hayashi, Dan Hartl, Andrew Haitt, Tom Hynes, Paula Jardieu, Karen Johnson, Dan Klessig, Jeff Kuret, Mike Laspia, Philip Leder, Fred Ledley, Vincent Marchesi, Rob Martienssen, Dusty Miller, Rick Myers, Karoly Nikolics, Luis Parada, Scott Putney, Don Rio, David Schlessinger, Matt Scott, John Sulston, Barbara Trask, Rebecca Ward, Robin Weiss, Jim Wells, Ted White, and Bob Williamson. Any errors, nevertheless, are ours, and not theirs.

Elizabeth Zayatz, the development editor, had the unenviable job of trying to make us concentrate on the work at hand when we wanted to be doing other things. She did wonderfully well keeping us at it and became a trusted friend. Bill O'Neal and Jodi Simpson, the manuscript editors, smoothed awkward passages and made us think more carefully about the information we were trying to convey. Janet Tannenbaum, the project editor, combed the manuscript and graphics with a thoroughness that must have required a magnifying glass. She guided us through a forest of edited manuscript, galleys, and page proofs. At times it seemed that never a day went by without an overnight package from Janet arriving on our desks. Alison Lew gave the text, illustrations, and cover their design and Bill Page looked after the art program—tasks that rapidly assumed epic proportions as we all worked to produce the figures that are an essential complement to the text. The figures were rendered by Network Graphics and Tomo Narashima. The beautiful and remarkable cover art is by Marvin Mattelson. Listening to him expound on surrealism and watching him sketch were highlights of the production. Julia De Rosa masterfully coordinated the production process to allow rapid completion of the project. Linda Chaput, President of Scientific American Books, was always patient and understanding. Her guidance and advice were invaluable when the going got tough, and our discussions with her were rewarding, and fun.

James D. Watson Michael Gilman
Jan Witkowski Mark Zoller
December, 1991

CHAPTER



In Vitro Mutagenesis

Recombinant DNA technology and DNA sequencing provided the tools to clone and characterize genes. As we learned in Chapter 8, simple inspection of gene sequences told us much about genomic organization. Functional sequences, such as transcriptional control elements, could often be identified by comparing sequences of a number of genes. However, to delve deeply into the structure and function of genes required the ability to change the DNA sequence and examine the effect of the change on gene function. For decades before the advent of recombinant DNA, this was done by classical genetics, the identification of mutant organisms with new properties. From the genetic properties of mutants, information about the structure and function of the underlying genes could often be inferred. This approach, however, was limited to organisms in which simple genetic analysis was possible—bacteria, yeast, fruit flies. Genetic analysis of more complex, longer-lived organisms like mice and men was slow and difficult.

Recombinant DNA changed all that. The ability to isolate genes as molecular clones, the development of tools to modify gene sequences in the test tube, and the power to return altered genes to the organism to test their function have revolutionized the way genetics is done in higher organisms. Because we now often work "backwards" from gene sequence to gene function, in contrast to

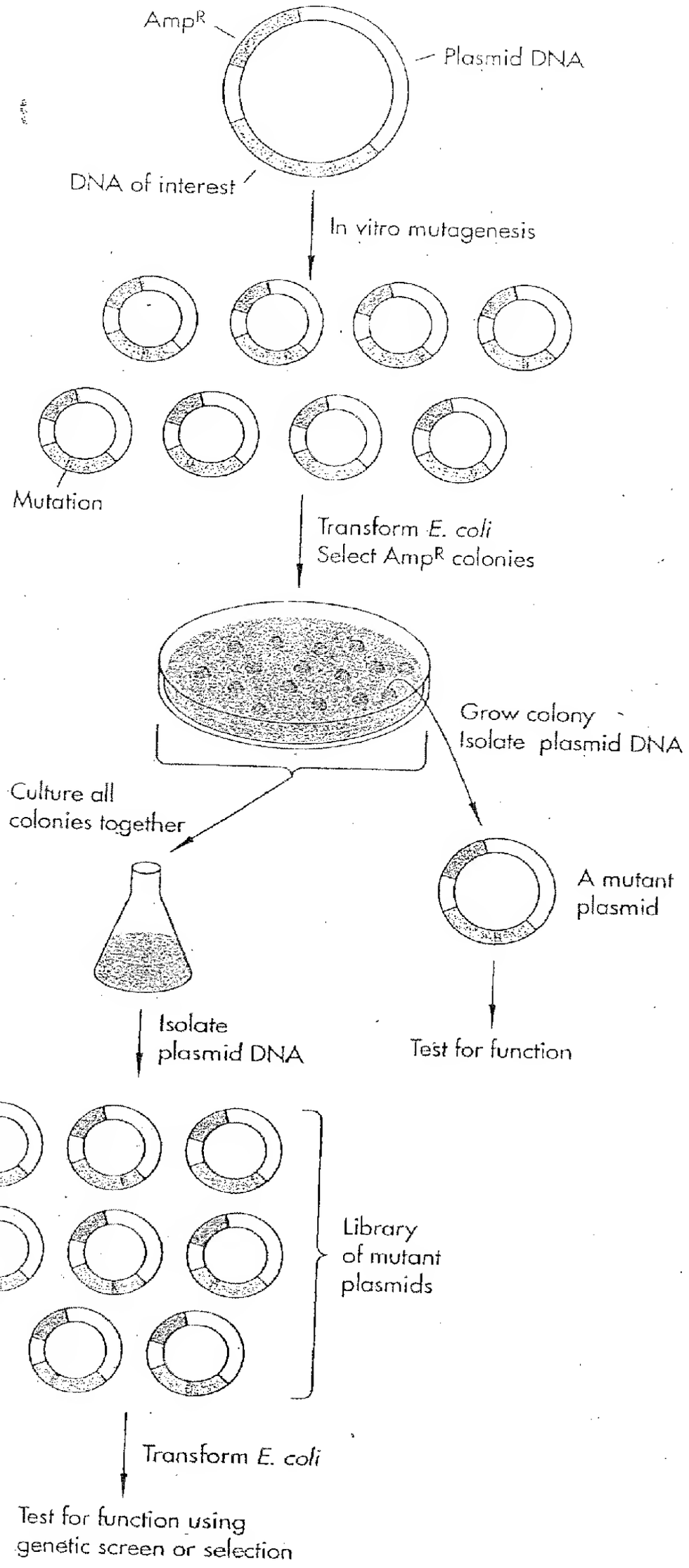


FIGURE 11-1
General strategy for an in vitro mutagenesis experiment. Most procedures for in vitro mutagenesis follow the same basic scheme: Plasmid DNA is "mutagenized" in vitro, then introduced into *E. coli* by transformation. Depending on the method, mutant clones can be isolated and tested individually, or a library of mutant plasmids can be obtained, which are tested using a genetic screen.

classical genetics, this new approach spawned by recombinant DNA is called *reverse genetics*. In this chapter we will learn ways to alter the sequence of a cloned gene at will and how these methods are used to understand the structure and function of genes and gene products.

In Vitro Mutagenesis Is Used to Study Gene Function

In vitro mutagenesis of cloned genes has become a standard tool in the functional analysis of nucleic acids and proteins. Most procedures follow the same basic scheme (Figure 11-1). Plasmid DNA containing the gene of interest is treated in vitro by some mutagenesis procedure that alters the DNA either chemically or enzymatically. The mutagenized plasmid DNA is introduced into *E. coli* by transformation, and colonies containing plasmid molecules are selected by antibiotic resistance. Mutants can be made one at a time, or hundreds of different mutants can be created in a single mutagenesis experiment. Mutant plasmids can be isolated from single colonies and tested individually. Alternatively, plasmid DNA can be prepared from pooled colonies and the resulting library tested en masse to identify mutant plasmids.

The various approaches to mutagenesis can be grouped broadly into random and site-directed methods. Random methods put mutations anywhere in a plasmid. They are best used to identify the location and boundaries of a particular function within a cloned DNA fragment and are most readily used for this purpose when a simple genetic screen (or selection) is available. A genetic screen or selection consists of a system to test the function of the DNA of interest in cells without having to isolate each plasmid individually. Random mutagenesis is often used as a first step, when little is known about the function encoded by particular DNA fragment. Analysis of random mutants generally provides only a simple identification of the functional region but does not explain how things work on a molecular level. The value of such a strategy is that it quickly helps to narrow down the focus of attention from a large DNA fragment to a smaller region that can be studied subsequently in greater detail. As we will learn, random mutagenesis can be accomplished by several different methods, such as altering the sequences within restriction en-

donuclease sites, inserting an oligonucleotide linker randomly into a plasmid, damaging plasmid DNA in vitro with chemicals, or incorporating incorrect nucleotides during in vitro DNA synthesis.

Once an important functional domain in a gene has been identified by random mutagenesis, site-directed methods—putting mutations precisely where they are needed—are used to define the role of specific sequences. In addition, directed mutagenesis provides a powerful tool for the analysis of protein function, by allowing researchers to make specific and subtle changes in the structure of the protein. A number of strategies have been developed to construct site-directed mutants in vitro, but this type of mutagenesis is best accomplished using synthetic oligonucleotides. With an oligonucleotide the desired sequence is simply built into the wild-type framework. Nowadays, oligonucleotide-directed mutagenesis reactions are relatively straightforward, and oligonucleotides are cheap and easy to obtain. The limitation of site-directed mutagenesis is that you must already have enough information to know what you wish to change. There are two standard ways of using oligonucleotides to construct site-directed mutants: mutagenesis by gene synthesis and mutagenesis by enzymatic extension of a mutagenic oligonucleotide. By using degenerate oligonucleotides (see Chapter 7) a set of "random" mutations at a specific site can also be made.

Restriction Endonuclease Sites Provide the Simplest Access for Mutagenesis

One of the first experiments done with a cloned DNA fragment is to map the positions of restriction endonuclease cleavage sites in the DNA by using a battery of different enzymes. Although this information could be precisely obtained from the DNA sequence, mapping restriction sites can be accomplished rapidly and is often done in conjunction with sequencing. Restriction endonuclease recognition sites provide the simplest way to modify a DNA clone in vitro (Figure 11-2). Cleaving plasmid DNA with a restriction enzyme that recognizes only one site produces a linear molecule. This serves as an entry point for modifying the DNA sequence in the vicinity of the restriction site. For example, the enzyme *Eco*RI recognizes the sequence GAATTC and produces ends with 5′ overhangs. The ends can be made even (blunt) by treating

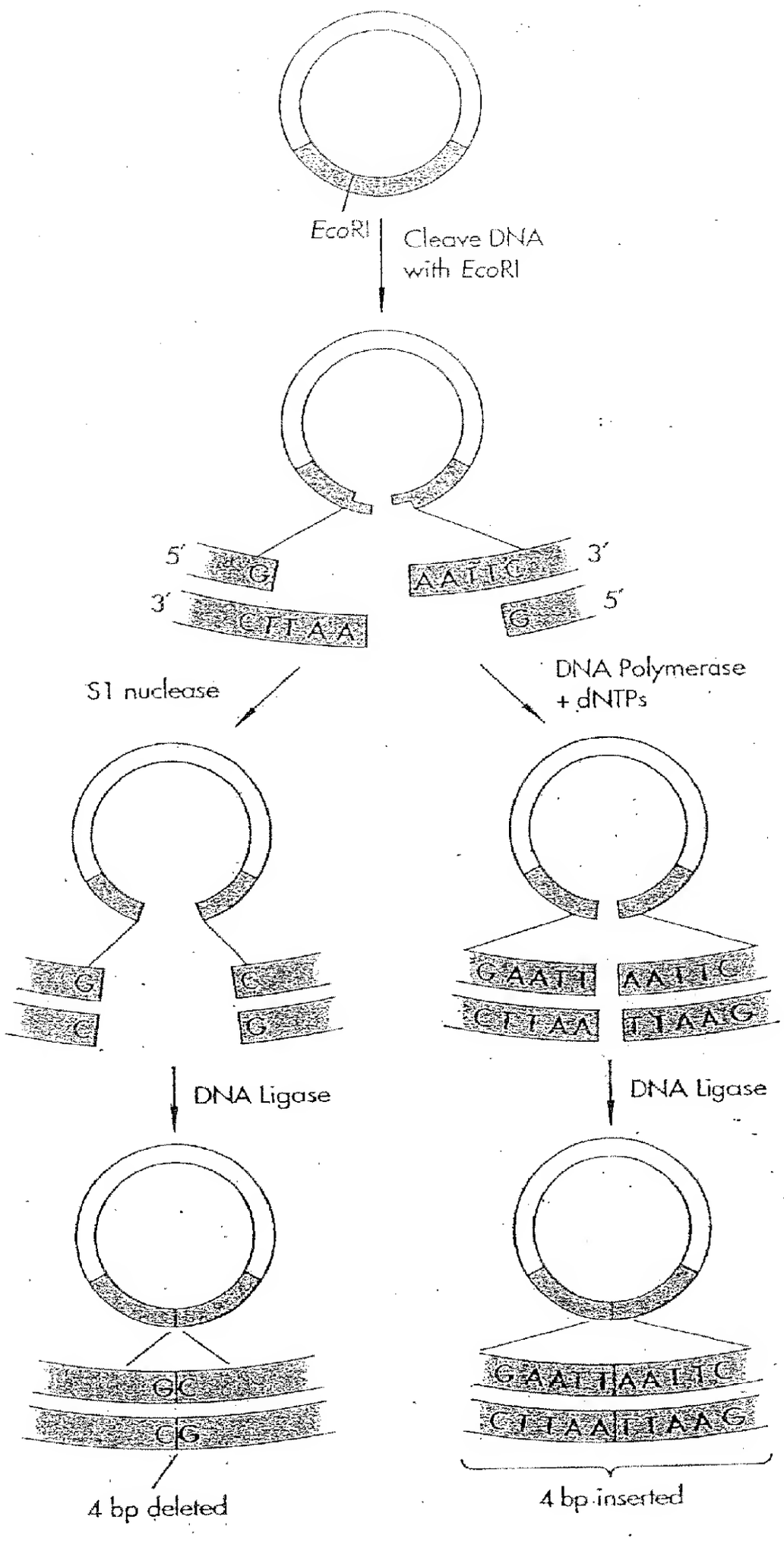
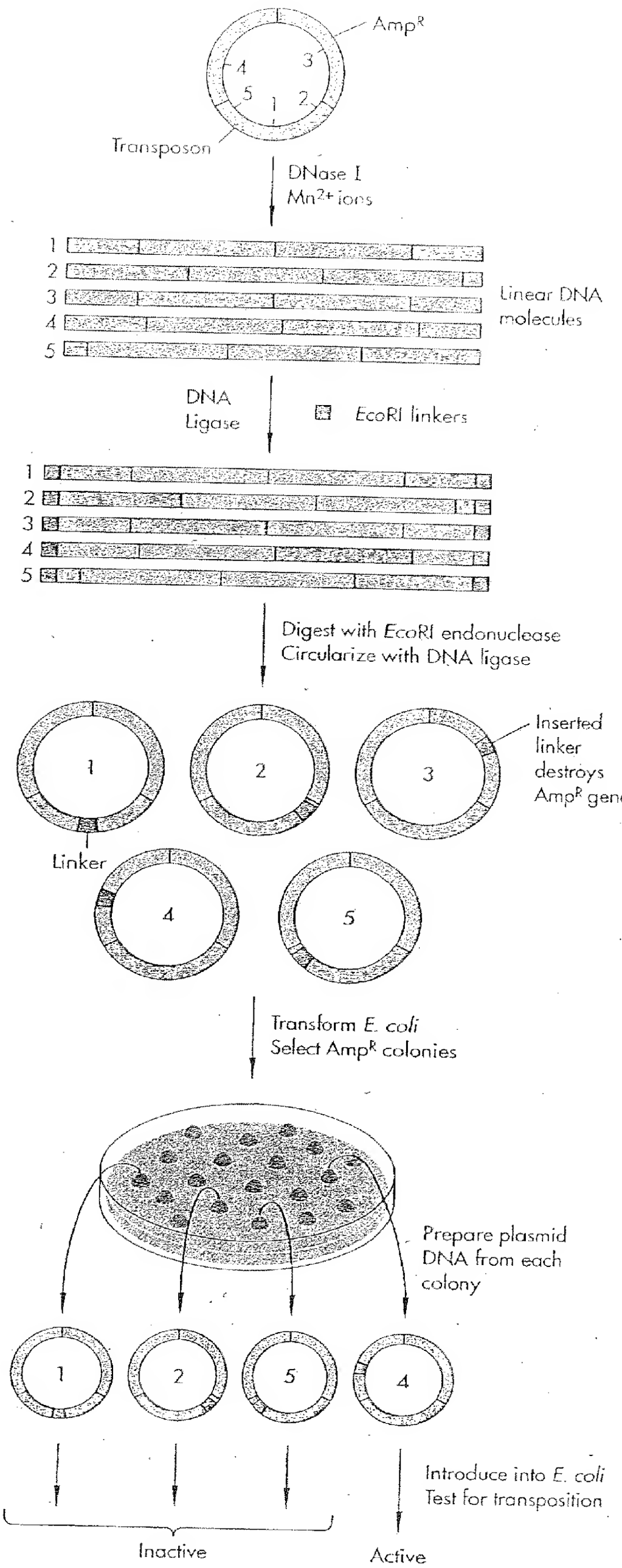


FIGURE 11-2
Creating a mutation by manipulation of a restriction site. Plasmid DNA is cleaved with *Eco*RI restriction endonuclease, which generates a linear fragment with 5′ ends that have four unpaired nucleotides (so-called sticky ends). Treatment with S1 nuclease (left) removes these nucleotides, and the linear fragment is then treated with DNA ligase. The resulting circular molecule contains a deletion of 4 bp. Alternatively, addition of DNA polymerase and deoxyribonucleotide triphosphates (dNTPs) to the plasmid cleaved by *Eco*RI extends the 3′ ends by DNA synthesis (right). After ligation, the resulting molecule contains an insertion of 4 bp. In both cases, the *Eco*RI site has been destroyed.

FIGURE 11-3
Linker insertion mutagenesis to map functional domains of a bacterial transposable element. The starting plasmid contains an intact transposon, an ampicillin-resistance gene for selection in *E. coli*, and sequences for plasmid replication. The DNA is treated with a low concentration of deoxyribonuclease I in the presence of Mn^{2+}. Under these conditions, the enzyme makes double-stranded cuts at random positions in the plasmid, generating a collection of linear DNA molecules broken at different positions. Oligonucleotide linkers encoding an *Eco*RI restriction site are added to the ends with DNA ligase, the linear molecules are treated with *Eco*RI endonuclease to create sticky ends on the linkers, and the molecules are recircularized. The circular molecules are transformed into *E. coli*, and ampicillin-resistant colonies are selected. Plasmid DNA is isolated from individual colonies, introduced into another strain of *E. coli*, and tested for activity of the transposon. The positions of the inserted linkers are mapped by restriction digestion. Linkers inserted in one region (blue) of the plasmid inactivated the transposon. No linker insertions in the ampicillin-resistance gene were recovered, because these plasmids would fail to yield a drug-resistant colony in the original selection of transformed *E. coli*.



the cleaved DNA with DNA polymerase in the presence deoxyribonucleotide triphosphates. The two blunt ends can then be linked together again (ligated) by incubating the linear plasmid molecule with DNA ligase. A few nanograms of DNA from the in vitro ligation reaction is used to transform *E. coli*, and the new modified plasmid is isolated from one of the resulting colonies. The net result of these manipulations is to insert 4 bp into the plasmid at the *Eco*RI site. Alternatively, a small deletion mutation can be made by treating the linearized DNA with S1 nuclease, which specifically digests single-stranded DNA. This creates blunt ends by removal of the four nucleotides that constitute the 5′ overhang generated by *Eco*RI at each end. Subsequent ligation of the DNA into a covalently closed circular molecule thus results in the deletion of 4 bp from the DNA. In each example, the new sequence no longer encodes the *Eco*RI recognition site. These types of manipulations, if done to a protein-coding sequence, would change the translational reading frame, resulting in production of a grossly altered protein. The major limitation of using restriction sites to make mutations is that there simply may not be sites in regions of the gene the experimenter wishes to alter.

Linker Insertion Is Used to Map a Bacterial Transposon

We have learned that it is a simple matter to cleave a plasmid with a restriction enzyme, blunt the ends by treatment with a DNA polymerase, and rejoin them by ligation. A variation on this technique is to rejoin the ends in the presence of a synthetic oligonucleotide "linker," often one that encodes a restriction site. Insertion of the linker disrupts the gene sequence; the position of the inserted linker can be easily mapped by cleavage of the plasmid with the restriction enzyme that cuts the linker.

A similar method was used to define the functional regions of a bacterial transposable element (a "jumping gene," see Chapter 10), by inserting linkers at many alternative positions throughout the element. To place linker insertions in the transposon, a plasmid carrying a clone of the transposon was treated with a nuclease that cleaved the plasmids at random positions (Figure 11-3). Cleavage conditions were adjusted so that each plasmid was cut just once on average. The linearized molecules were isolated and ligated into circles again in the presence of an 8-bp linker oligonucleotide containing an *Eco*RI restriction site, resulting in insertion of the linkers into random sites, one in each plasmid. The resulting plasmids were transformed into *E. coli* and, using a genetic screen, examined to see if the transposon could jump. Insertion of a linker into a region of the transposon critical for its function inactivates it, presumably by putting a protein-coding sequence out of frame. By mapping the positions of the inserted linkers by restriction analysis, the locations of functional regions of the transposon were deduced.

Construction of Nested Deletions Maps the Boundaries of a Transcriptional Control Region

Transcription of the gene encoding the 5S ribosomal RNA molecule is carried out by RNA polymerase III (pol III, see Chapter 8). To identify the sequences within the 5S gene required for transcription by pol III, a series of deletion mutations was made and tested for their ability to support accurate transcription. Two sets of deletions were made. One was made by cutting a plasmid carrying a cloned 5S gene at a restriction site on the 5′ side of the gene. The linearized plasmid was treated with a combination of nucleases that digested away DNA from the ends of the molecule (Figure 11-4). The amount of DNA removed was controlled by varying the time, temperature, or enzyme concentration in the reaction. A second set of deletions was generated from plasmid DNA cleaved at a site on the 3′ side of the gene. The result was two sets of plasmids with progressively larger deletions toward the gene from both directions. Testing these genes revealed that only deletions entering a 35-bp region within the transcribed region of the 5S gene abolished transcription by pol III. Therefore, this deletion analysis mapped the transcriptional regulatory element to this 35-bp stretch, which has subsequently been analyzed in much greater detail by site-directed mutagenesis.

Several different types of enzymes can be used to produce deletions. Generally, these enzymes delete DNA from both ends of a linearized plasmid molecule. Often, however, one end of the molecule contains sequences that need to be retained in the plasmid because, for example, they are required for plasmid replication. In the 5S gene deletion experiment, this limitation was accommodated by isolating the deleted gene fragments and recloning them into a new vector. Alternatively, a strategy can be used that limits deletion to one end of a linearized plasmid molecule (Figure 11-5). This method is widely used to generate nested deletions for DNA sequencing (see Chapter 7).

Linker-Scanning Mutagenesis Permits Systematic Analysis of Promoters

Deletion mutagenesis of the 5S gene mapped the boundaries of the transcriptional control region in the gene. But not all the nucleotides within the boundaries of that 35-bp region are necessarily critical for function. Therefore, methods were needed to change individual nucleotides in a target without generating gross deletions or other rearrangements. This was accomplished for a viral promoter using an elegant ad-

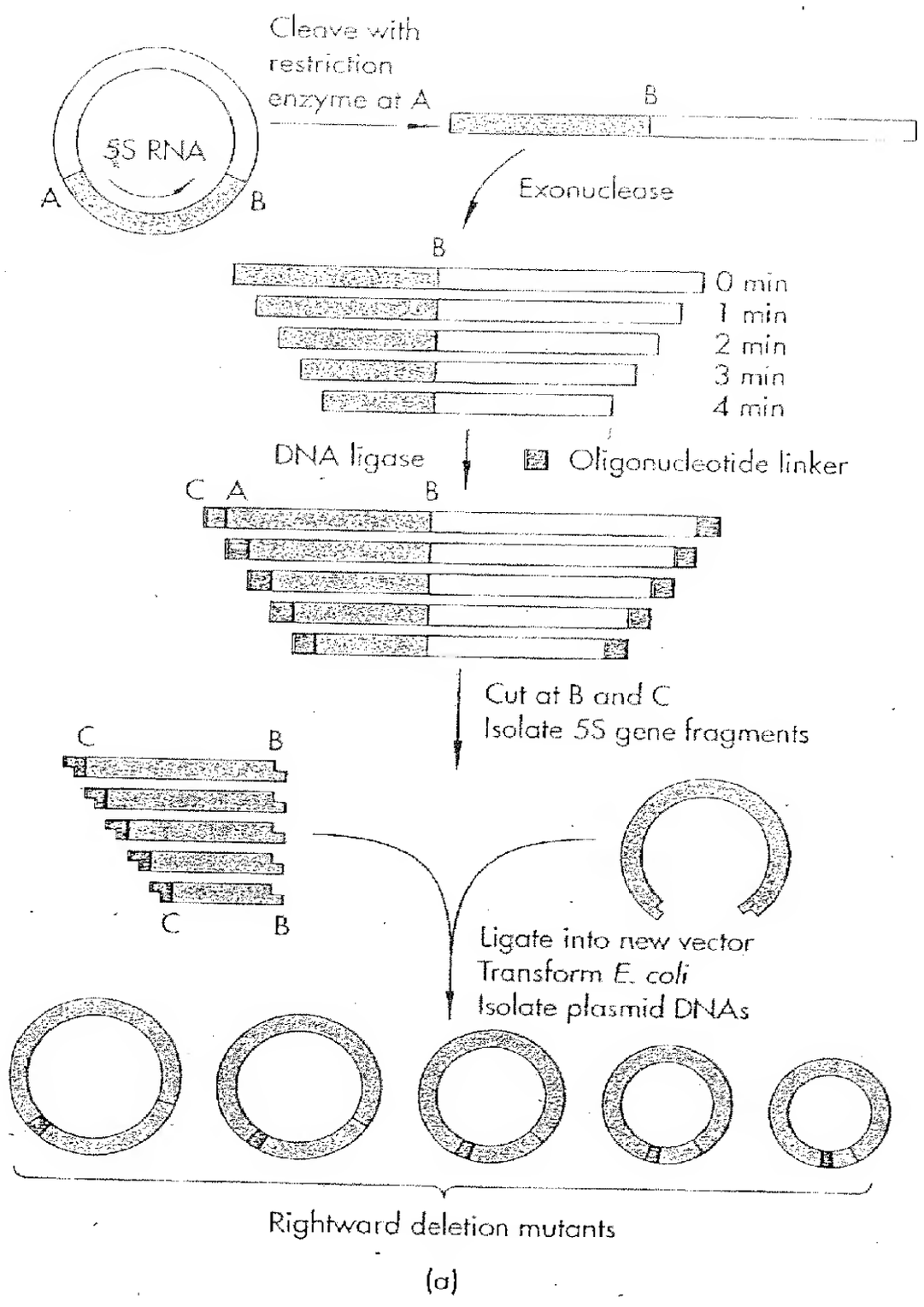
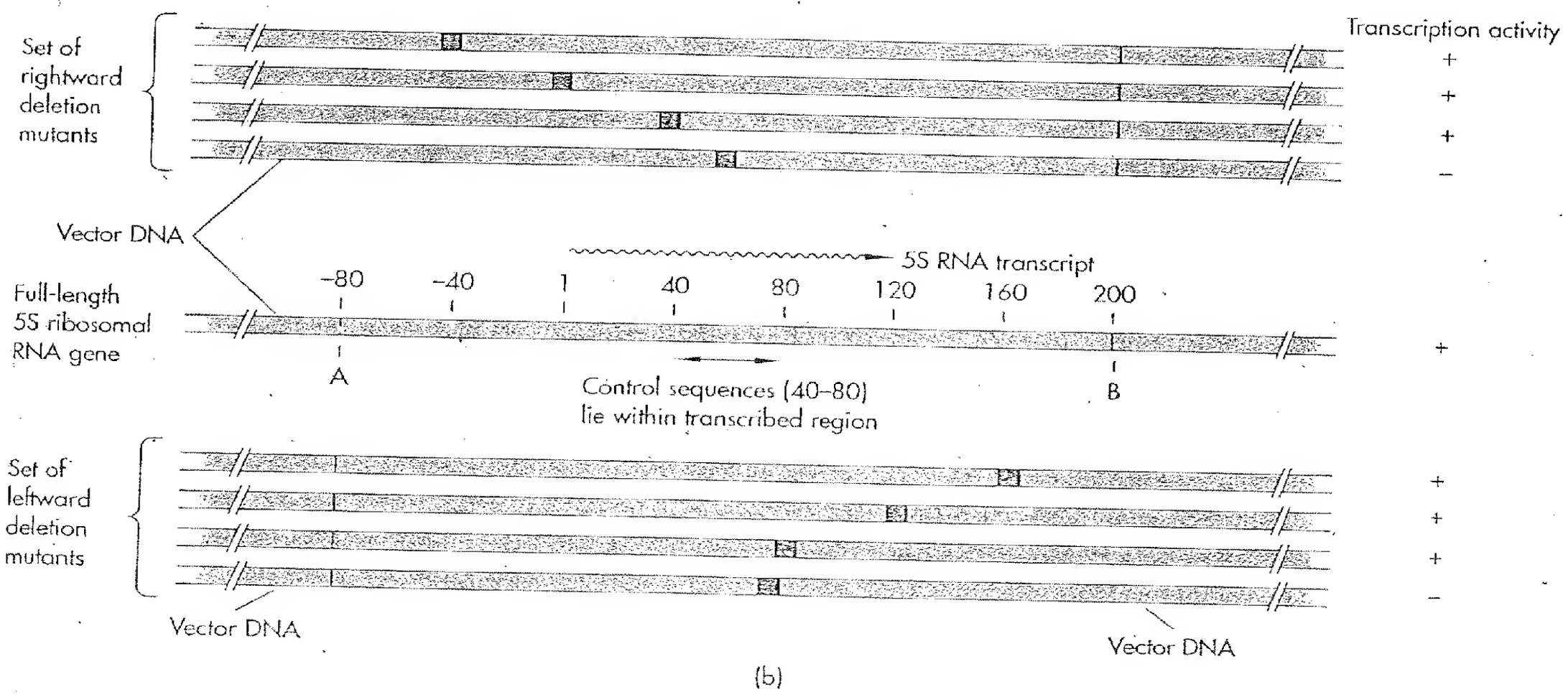


FIGURE 11-4
Construction of a nested set of deletion mutants to map the transcription control region of a 5S ribosomal RNA gene. (a) A plasmid clone was linearized with a restriction enzyme at a position (A) on the 5′ side of the gene. The linear fragments were treated with an exonuclease, which digests DNA from both ends of the molecule. Portions of the reaction were removed at different times to recover populations of molecules with progressively larger deletions. Linkers were added to the ends, and the molecules were cleaved with restriction enzymes specific for sites B and C to separate the 5S gene fragments from the remnants of the vector. The fragments were recloned into a new vector, generating the set of rightward deletion mutants. To create the leftward deletion mutants, this process was repeated after cleaving the plasmid at restriction site B. (b) Individual plasmids were isolated after transformation, their deletion endpoints determined by DNA sequencing, and their ability to support transcription by RNA polymerase III tested with an in vitro assay. As can be seen by comparing transcription activity with the extent of deletion, transcription is inhibited when the rightward (5′) deletions enter the +40 region and when the leftward (3′) deletions pass the +80 point. This suggests that the transcription control region lies between +40 and +80.



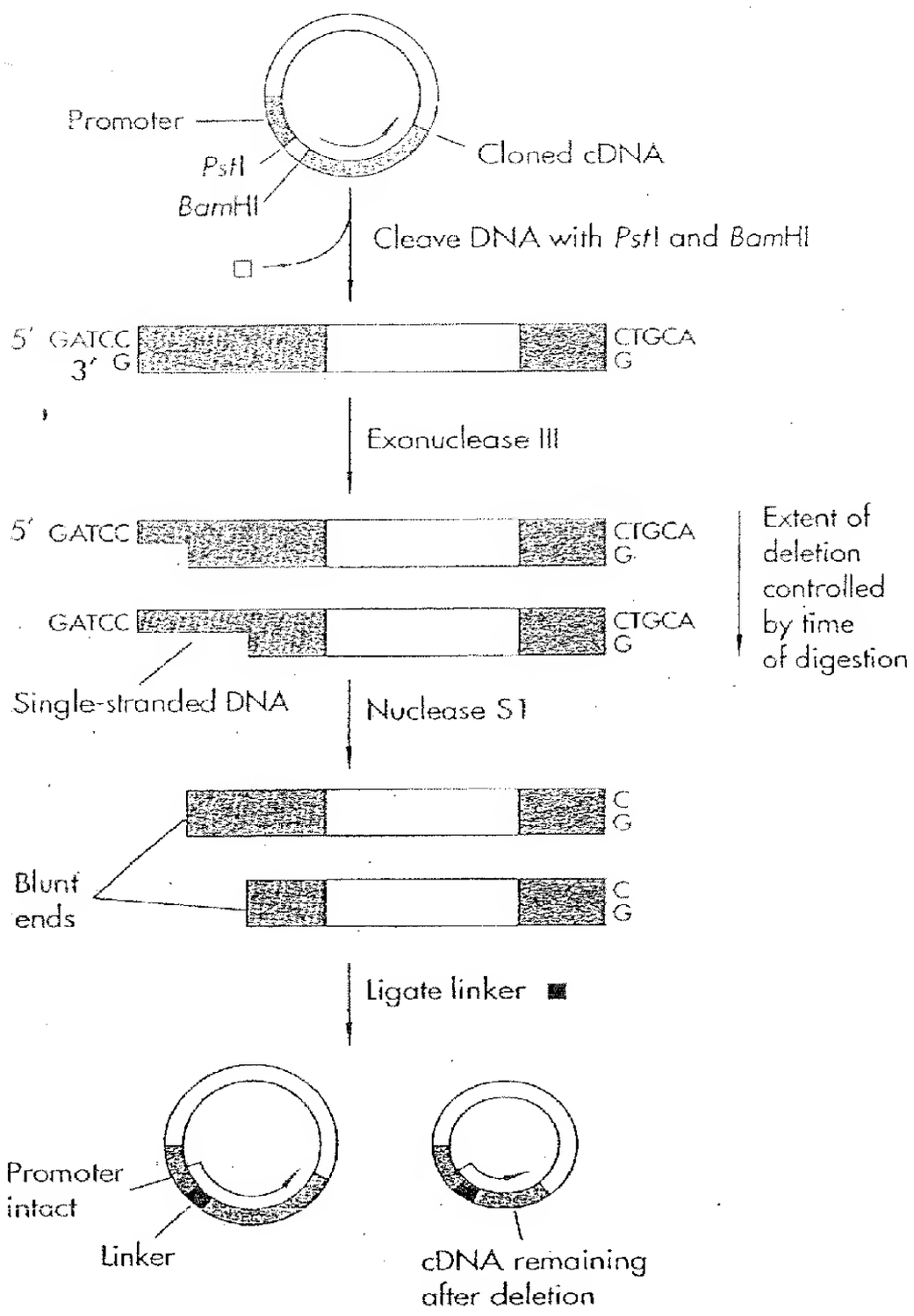


FIGURE 11-5
Construction of unidirectional deletions using exonuclease III. Exonuclease III attacks preferentially the 3′ end of a linear DNA molecule with 5′ protruding nucleotides. Therefore, by cleaving a plasmid molecule at adjacent sites with *Bam*HI, which leaves a 5′ overhang, and *Pst*I, which leaves a 3′ overhang, only the end generated by *Bam*HI is attacked by exonuclease III. After exonuclease III treatment, the remaining single-stranded tail (along with the overhang at the other end) is removed with S1 nuclease, which digests only single-stranded DNA. An oligonucleotide linker is attached, and the fragments are ligated to form closed circular molecules. In the experiment shown here, deletions are being used to map the functional domains of a cloned gene inserted in an expression vector. This strategy allows deletions to be made only in the cloned gene, without damaging the promoter sequence.

aptation of deletion mutagenesis called *linker scanning*. Using the methods outlined in Figure 11-4, two sets of plasmids were constructed that contained deletions within the promoter. One set of deletions started from a site beyond the 5′ end and proceeded toward the gene, leaving the 3′ end intact; the other set started at a point within the gene and proceeded in the opposite direction, leaving the 5′ end intact. Each deletion terminated with a 10-bp *Bam*HI linker. The extent of the deletion in the DNA was determined for each plasmid by DNA sequencing. Pairs of plasmids from the two deletion sets with endpoints 10 bp apart were recombined at their *Bam*HI sites (Figure 11-6). The effect was to preserve the length and organization of the promoter—thought to be important for promoter function—but to replace various 10-bp segments of wild-type promoter sequence with the sequence in the linker. Thus, this experiment created a library of promoter mutants of similar structure but with nucleotide substitutions clustered within 10-bp windows located at various sites in the promoter. This collection of mutants spanned the length of the promoter. The results of this analysis were discussed in Chapter 9. At the time, this experiment represented the most thorough analysis of a promoter in a mammalian gene.

Random Nucleotide Substitutions Are Obtained by Chemical Modification of DNA or by Enzymatic Misincorporation

While linker scanning allows the creation of nucleotide substitutions, each mutant generally contains several substitutions, and the positions of the mutations depend on the availability of appropriately placed deletions. Therefore, several strategies have been developed for placing *single* nucleotide substitutions at random positions in a DNA molecule. The simplest methods employ chemicals that modify or damage DNA. Generally, plasmid DNA or DNA fragments are treated with chemicals, transformed into *E. coli*, and propagated as a library of mutant plasmids. Chemicals most commonly used for in vitro mutagenesis include sodium bisulfite, which deaminates cytosine residues to uracil, and reagents that damage or remove

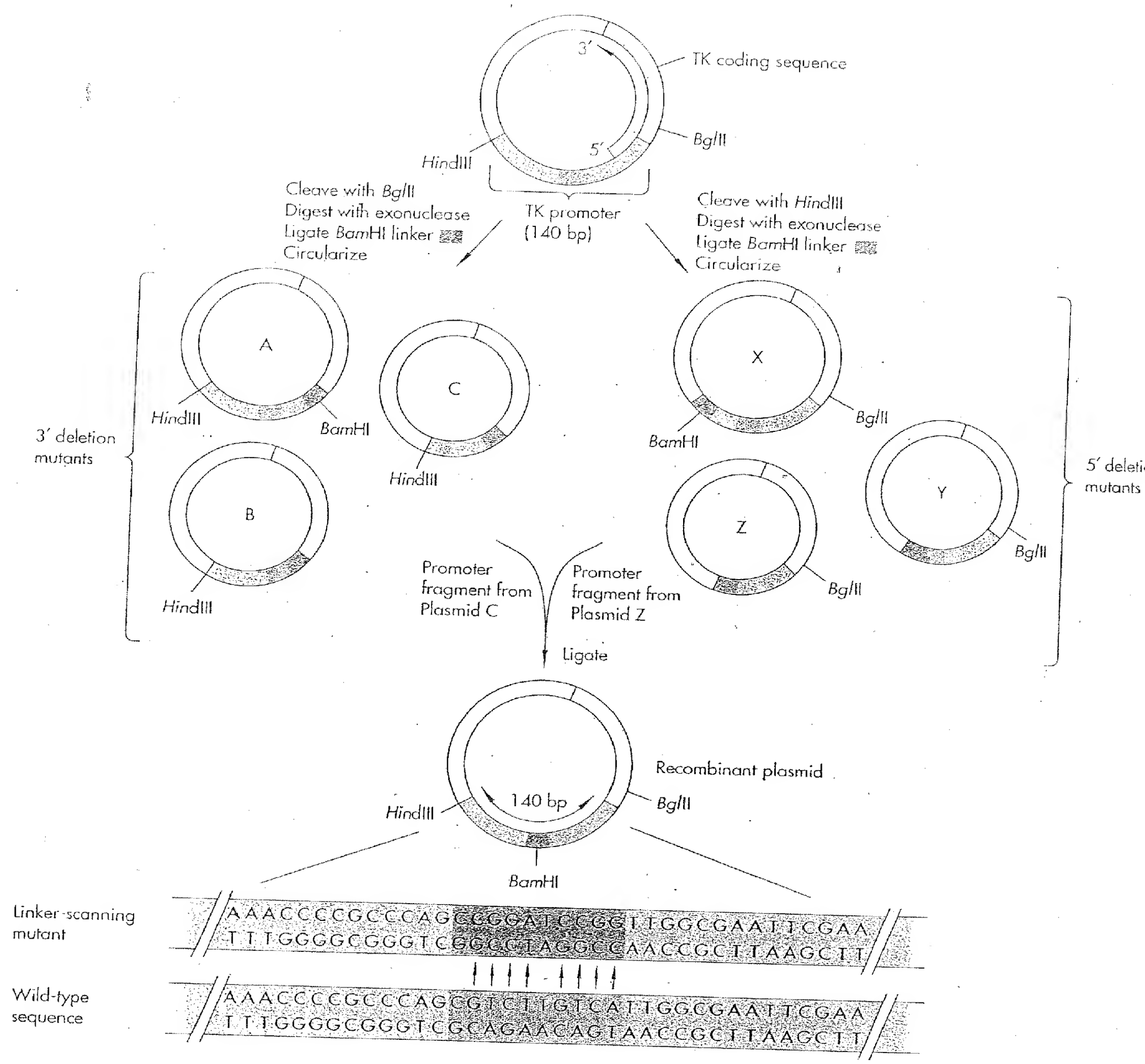


FIGURE 11-6
Linker-scanning mutagenesis of the viral promoter for the thymidine kinase (TK) gene. Two sets of deletion mutants were made beginning from restriction sites on the 5′ and 3′ sides of the promoter by the method described in Figure 11-4. (The promoter is divided into three colors to make the extent of deletion more obvious.) Approximately one hundred plasmids were sequenced to determine their deletion endpoints. Pairs of deletion fragments, where the 5′ deletion of one fragment ended precisely 10 bp downstream from the endpoint of the 3′ deletion of the other fragment, were identified. The *Hin*dIII–*Bam*HI fragment of the 3′ deletion mutant and the *Bam*HI–*Bgl*II fragment of the 5′ deletion mutant were joined via their *Bam*HI sticky ends and cloned into a new plasmid. This strategy yields molecules like the one shown at the bottom: they are wild-type in sequence except for the substitution of the 10-bp *Bam*HI linker in place of the sequence between the two deletion endpoints. In the example shown, this results in a cluster of eight nucleotide substitutions (arrows).

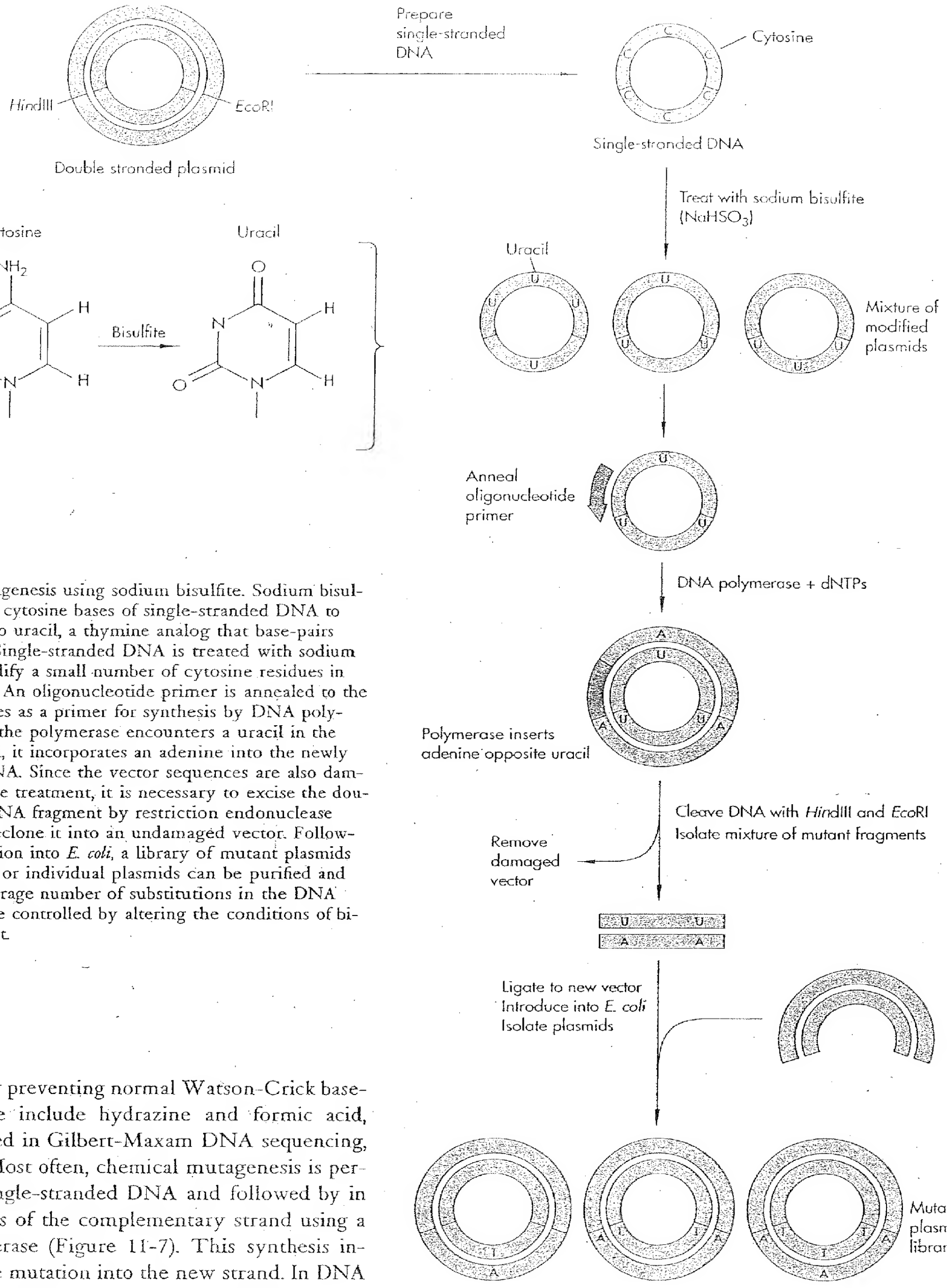


FIGURE 11-7
Chemical mutagenesis using sodium bisulfite. Sodium bisulfite reacts with cytosine bases of single-stranded DNA to convert them to uracil, a thymine analog that base-pairs with adenine. Single-stranded DNA is treated with sodium bisulfite to modify a small number of cytosine residues in each molecule. An oligonucleotide primer is annealed to the DNA and serves as a primer for synthesis by DNA polymerase. When the polymerase encounters a uracil in the template strand, it incorporates an adenine into the newly synthesized DNA. Since the vector sequences are also damaged by bisulfite treatment, it is necessary to excise the double-stranded DNA fragment by restriction endonuclease cleavage and reclone it into an undamaged vector. Following transformation into *E. coli*, a library of mutant plasmids can be isolated or individual plasmids can be purified and tested. The average number of substitutions in the DNA fragment can be controlled by altering the conditions of bisulfite treatment.

bases, thereby preventing normal Watson-Crick base-pairing (these include hydrazine and formic acid, which are used in Gilbert-Maxam DNA sequencing, Chapter 5). Most often, chemical mutagenesis is performed on single-stranded DNA and followed by in vitro synthesis of the complementary strand using a DNA polymerase (Figure 11-7). This synthesis incorporates the mutation into the new strand. In DNA treated with bisulfite, an adenine nucleotide is incorporated opposite the uracil; after transformation into

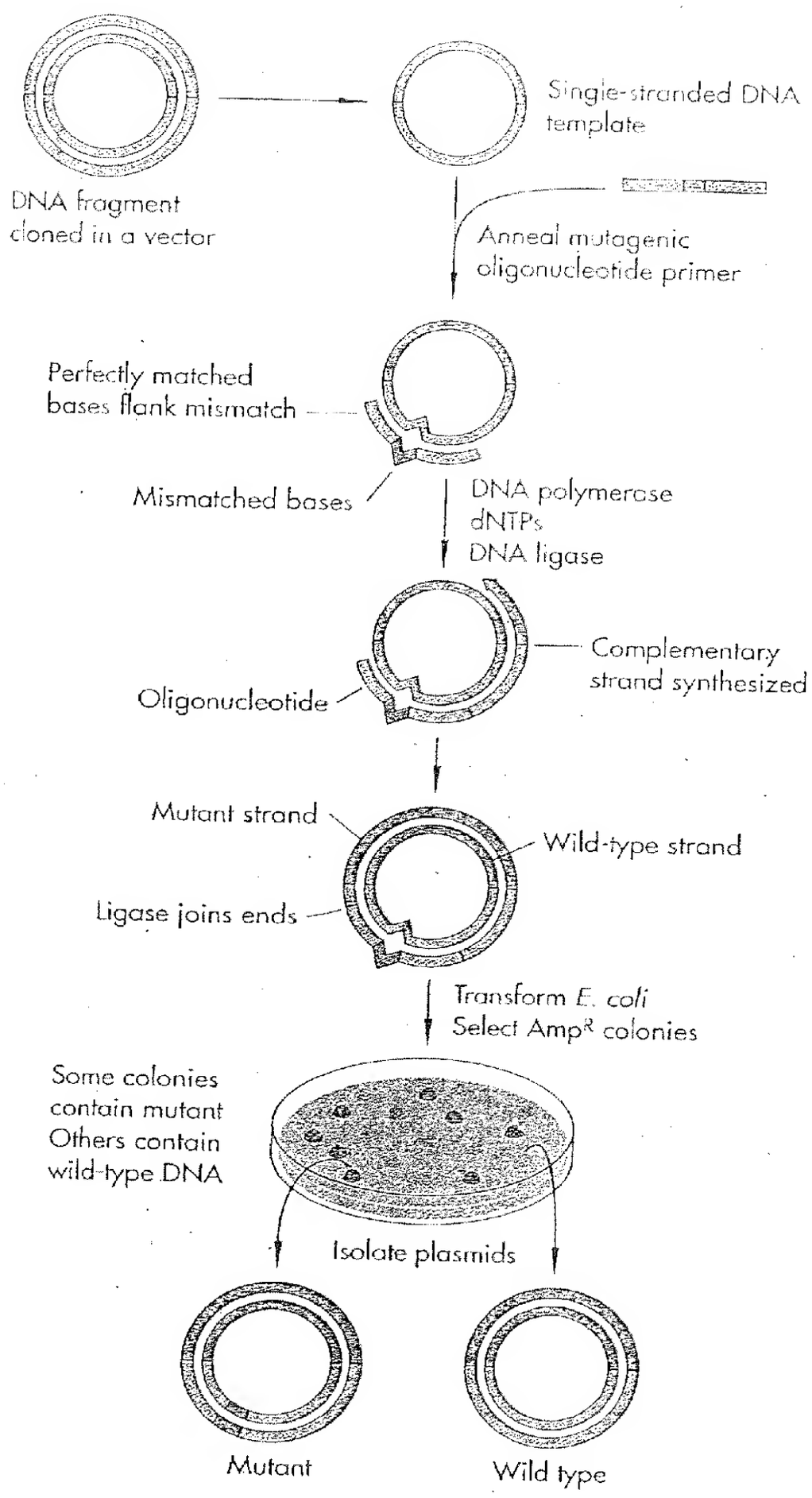


FIGURE 11-8
Oligonucleotide-directed mutagenesis by enzymatic primer extension. A "mutagenic" oligonucleotide encoding the desired mutation embedded in wild-type flanking sequence is annealed to a single-stranded DNA template. The sequence of the oligonucleotide is complementary to the template except for the nucleotides that define the mutation. Generally, the mutagenic oligomer is designed so that the mismatched nucleotides are positioned in the middle and there are at least 8 to 12 nucleotides on either side that base-pair with the template DNA. The mutagenic oligonucleotide serves as a primer for DNA synthesis by DNA polymerase. Once the entire template has been copied, the ends of the newly synthesized strand are covalently linked by DNA ligase. The heteroduplex DNA is transformed into *E. coli*. Theoretically, both strands can replicate, segregating into separate mutant and wild-type plasmids. In practice, however, most colonies contain only one or the other, because enzymes in the cell recognize and repair mismatched nucleotides in the heteroduplex before replication. Plasmid DNA is isolated from the resulting colonies and is screened to identify mutants.

E. coli, the wild-type C·G base pair becomes a T·A pair. In DNA treated with reagents that eliminate bases, any nucleotide can be incorporated opposite the "abasic" site, which still retains its deoxyribose backbone although it has lost its base. The major limitation of chemical mutagenesis is the specificity of the individual reagents: bisulfite mutagenesis, for example, changes only cytosines.

All possible nucleotide substitutions can be generated using enzymatic misincorporation. Here the strategy is to perform in vitro DNA synthesis under nonideal conditions—suboptimal ionic conditions, unbalanced concentrations of nucleotide precursors—that encourage DNA polymerase occasionally to incorporate the wrong nucleotide during synthesis. For example, synthesis is carried out in the presence of high concentrations of three of the precursors and a very low concentration of the fourth. At positions that normally call for the fourth (scarce) nucleotide, one of the others is sometimes incorporated instead. These methods also exploit DNA polymerases that lack a proofreading activity—a 3′ to 5′ exonuclease mechanism that checks each base pair after incorporation and removes nucleotides that are mismatched. *Thermus aquaticus (Taq)* DNA polymerase, used in the polymerase chain reaction (Chapter 6), lacks such an activity. Though this is a problem when accuracy of synthesis is required, the PCR is a very simple and efficient way to introduce random nucleotide substitutions into a DNA fragment.

A general problem with random mutagenesis approaches is that they often produce mutants with more than one substitution. Multiple substitutions in a single mutant complicate the interpretation of an experiment, because it isn't clear which substitution (or which combination of substitutions) is responsible for

observed changes in the properties of the mutant. Extraordinary methods have been used to circumvent this problem—essentially, significantly reducing the extent of mutagenesis and using enrichment protocols to find rare mutants—but almost all these procedures have been supplanted by new methods that use synthetic oligonucleotides.

Synthetic Oligonucleotides Facilitate Mutagenesis

Most of the methods for mutagenesis we have discussed so far have some significant shortcoming—they rely on fortuitous access to a sequence via a restriction site, forced entry through deletion strategies, or tedious screens to find randomly generated mutations in the region of interest. To be most powerful, mutagenesis must allow the experimenter to place *any* modification at *any* position desired in cloned DNA. This has become not only possible, but simple and cheap, with the advent of synthetic DNA oligonucleotides. Oligonucleotides provide the means to design a particular mutation and then to place it precisely where you want it.

The simplest method for doing oligonucleotide-directed mutagenesis is by enzymatic primer extension (Figure 11-8). In this method, an oligonucleotide is designed that carries the mutation flanked by 10 to 15 nucleotides of wild-type sequence. This "mutagenic" oligonucleotide is hybridized to its complementary sequence in single-stranded wild-type DNA prepared from a phage or phagemid clone, forming a heteroduplex with mismatched nucleotides at the site of the mutation. Although the oligonucleotide is not perfectly complementary, it will anneal if the hybridization conditions are not very stringent. The oligonucleotide serves as a primer for in vitro enzymatic DNA synthesis by a DNA polymerase that converts the single-stranded DNA into double-stranded form, using the wild-type strand as template. In this way, all regions of the plasmid except the region containing the mutagenic oligonucleotide will be wild-type in sequence. Once the primer has been extended completely around the template, the ends of the newly synthesized strand are ligated, forming a double-stranded circular DNA molecule. This heteroduplex DNA—one strand has the wild-type sequence and the other strand has the mutant sequence—is transformed into *E. coli*, where either strand can be replicated. By the time a colony grows up, however, it usually contains only one type of plasmid, wild-type or mutant. The types of mutations that can be made by this approach range from single nucleotide substitutions to deletions or insertions, limited only by the size of the oligonucleotide needed.

Mutant Clones Can Be Identified by Hybridization and DNA Sequencing

Theoretically, half the daughter molecules of a mutagenesis reaction will be wild-type and half mutant. In practice, however, the precentage of mutant plasmids is often much lower. This is due to a variety of technical factors, but the consequence is that methods for identifying or enriching mutant clones are vital. Mutant molecules can be distinguished from wild-type if there is gain or loss of a restriction site. Alternatively, the oligonucleotide that was originally used to make the mutation can be used as a hybridization probe to distinguish mutant from wild-type molecules (Figure 11-9). The mutagenic oligonucleotide is radioactively labeled with ^{32}P-ATP and hybridized to DNA from bacterial colonies on nitrocellulose filters, as described in Chapter 7. If the temperature of the hybridization is raised in 5 or 10°C increments, a point can usually be reached at which the labeled oligonucleotide will hybridize only to the mutant molecules (to which it is perfectly complementary) and not to the wild-type molecules, because the hybrid is destabilized by the mismatched nucleotides. Plasmid DNA is isolated from an *E. coli* colony that strongly hybridizes to the probe. Verification that the desired mutation was made is accomplished by sequencing the DNA of this putative mutant clone. This technique can identify one mutant clone among several hundred wild-type clones.

Several clever methods enrich for mutant clones so that the tedious task of screening by hybridization is not necessary. In one of these techniques, the template DNA is biologically marked so that it is destroyed after transformation into *E. coli* and the mutant strand

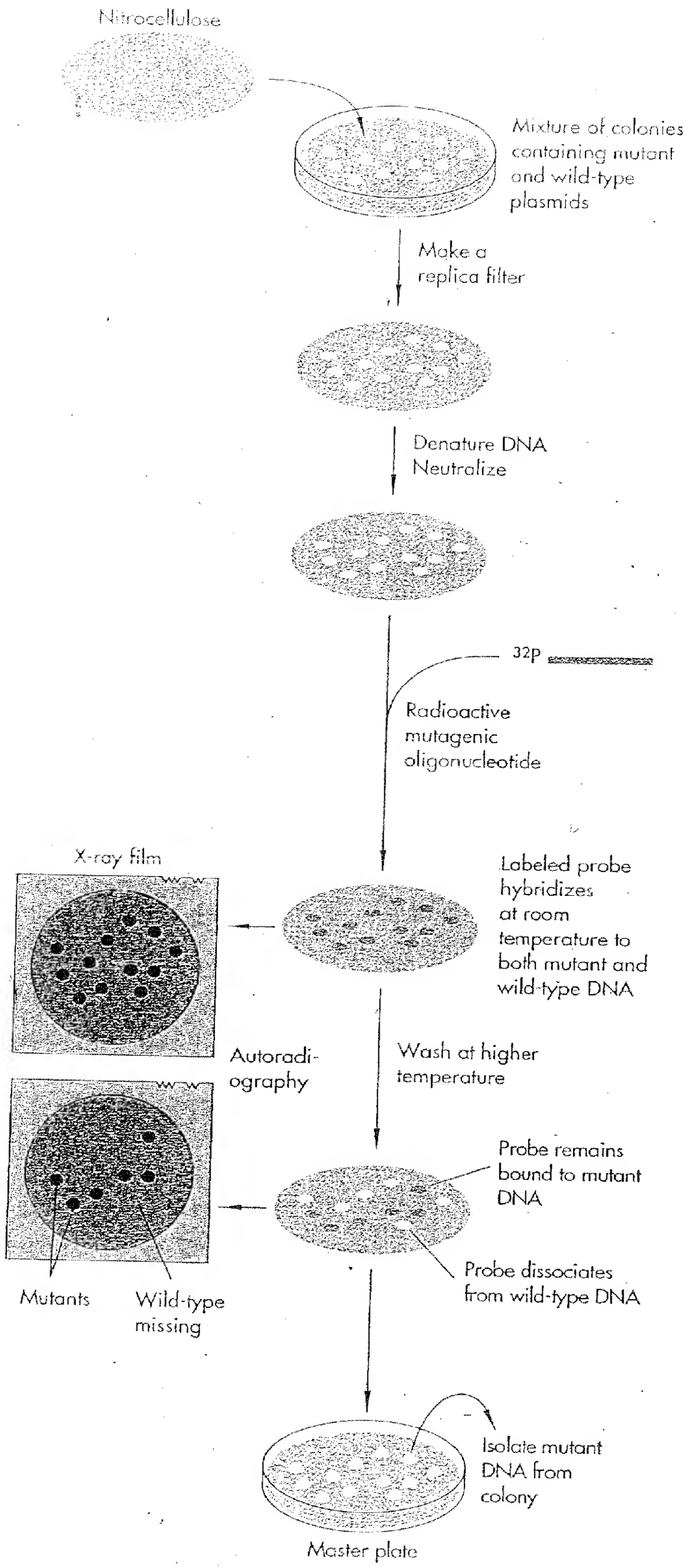


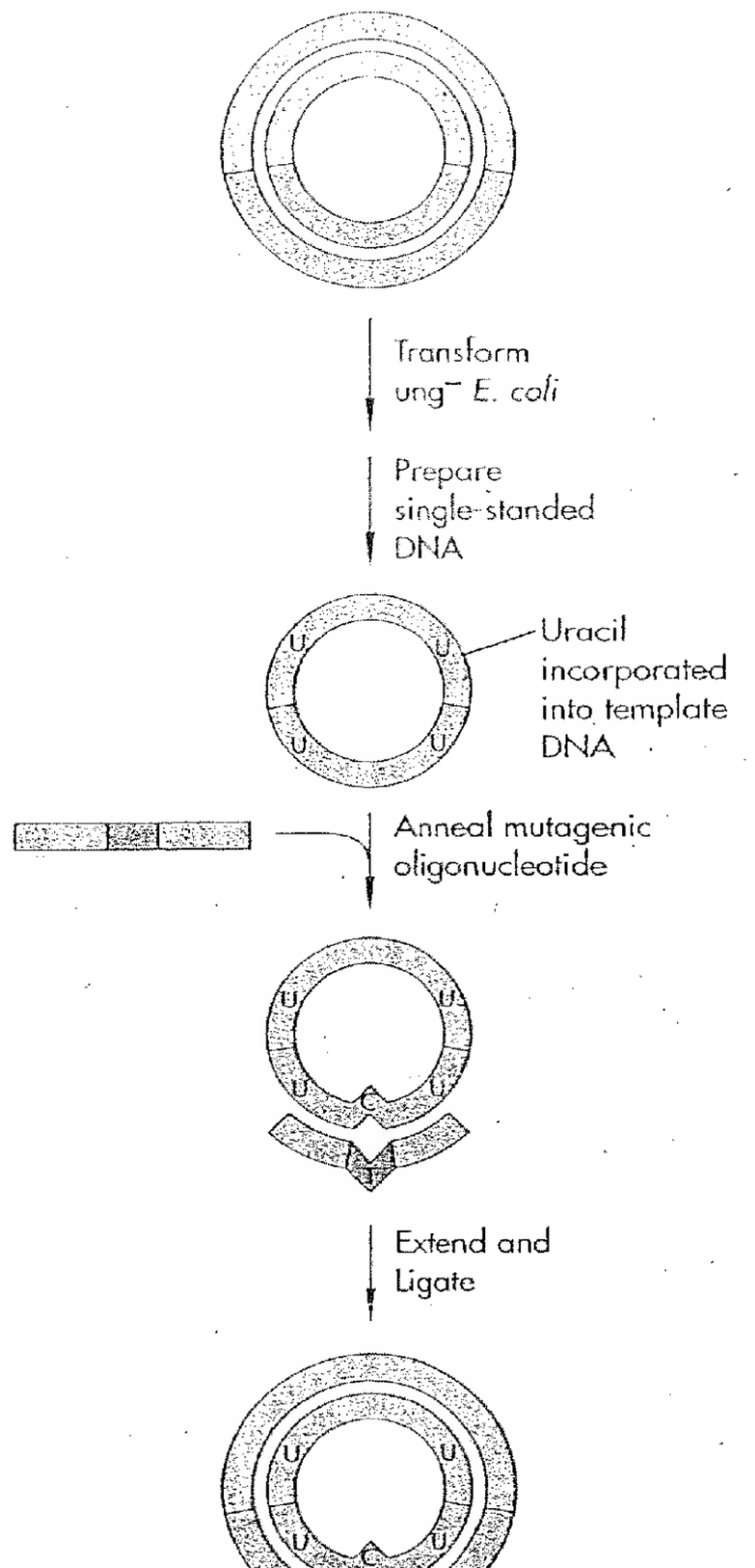
FIGURE 11-9
Searching for mutant plasmids using the mutagenic oligonucleotide as a probe. Colonies (or plaques) resulting from transformation by mutagenized plasmids (see Figure 11-8) are prepared for colony hybridization on nitrocellulose filters using methods described in Chapter 7. The mutagenic oligonucleotide is radioactively labeled by phosphorylating its 5′ end using ^{32}P-ATP and polynucleotide kinase. The labeled oligonucleotide is hybridized to the plasmid DNA on the nitrocellulose filters. At low temperature, the oligonucleotide will hybridize to both mutant and wild-type DNAs. As the temperature is increased, the mismatched oligonucleotide hybridized to the wild-type plasmid DNA begins to dissociate from the wild-type clones. Eventually a temperature is reached at which the mismatched oligomers completely dissociate from the wild-type clones but remain hybridized to the mutants. Since the oligonucleotide is radioactively labeled, the nitrocellulose filter is exposed to x-ray film and mutant clones are identified by the presence of a strong signal on the autoradiograph. Mutant plasmid DNA is then isolated from the corresponding colony on the master plate, using the replica filter as a guide.

is preferentially replicated (Figure 11-10). In a second method, the template strand is enzymatically destroyed before transformation. Both methods can yield mutants at a frequency of greater than 50 percent, so that plasmid DNA is simply isolated from three or four randomly picked colonies and analyzed by DNA sequencing with the expectation that a mutant will be found among the DNA selected.

Oligonucleotide Cassettes Provide a Simple Method for Introducing Directed Mutations

We learned earlier that restriction enzyme sites provide access to a cloned DNA for mutagenesis. If two restriction sites are close together, the intervening fragment can be removed and replaced with a synthetic double-stranded fragment (a *cassette*) made from two complementary single-stranded oligonucleotides carrying any desired sequence. Often, however, convenient restriction sites are not available; fortunately, it is a simple matter to create them using the oligonucleotide-directed mutagenesis procedures described in the previous sections. Once the sites are in place, any

FIGURE 11-10
Enrichment for oligonucleotide-directed mutants by using a uracil-containing template. Single-stranded template DNA is prepared in a strain of *E. coli* that lacks the enzyme uracil deglycosidase (ung$^-$), so that it contains several uracil residues in place of thymines. (Although uracil is not usually incorporated into DNA, it is not actually mutagenic and it does form a base-pair with adenine.) The mutagenic oligonucleotide is annealed and primes the synthesis of a strand that extends around the template in a reaction using the four standard dNTPs (as in Figure 11-8). Following ligation, the heteroduplex DNA molecules are introduced into an ung$^+$ strain of *E. coli*. Once in the cell, the wild-type (template) strand is attacked by uracil deglycosidase, which causes breaks in the DNA strand, and the DNA strand is degraded before it can be replicated. Since the strand containing the mutagenic oligonucleotide does not contain uracil, it is not attacked and is replicated normally. When this procedure is used, 50 percent or more of colonies contain mutant plasmids.



number of new mutants can be made by inserting synthetic fragments into the plasmid (Figure 11-11), just as different cassettes can be inserted into a tape player.

This method of cassette mutagenesis was the basis for an elegant experiment that verified a structural model for DNA recognition by phage repressors. The repressors of the λ-like phages 434 and P22 contain a helix-turn-helix structure (see Chapter 9) that recognizes the operator DNA in the phage genome. It was hypothesized that amino acid side chains on one face of an α helix in the repressor protein make sequence-specific contacts with operator DNA. To test this hypothesis, a *helix swap* was performed (Figure 11-12). Oligonucleotides were synthesized that encoded the amino acids of the helix in the 434 repressor, with the five positions thought to contact DNA changed to those found in the P22 repressor. This synthetic fragment was swapped for the natural fragment in the 434 gene. The resulting hybrid protein gained the recognition specificity of the P22 repressor, demonstrating that this helix indeed contacts the DNA.

Cassette mutagenesis with degenerate oligonucleotides can be used to create a large collection of random mutations in a single experiment. This method was

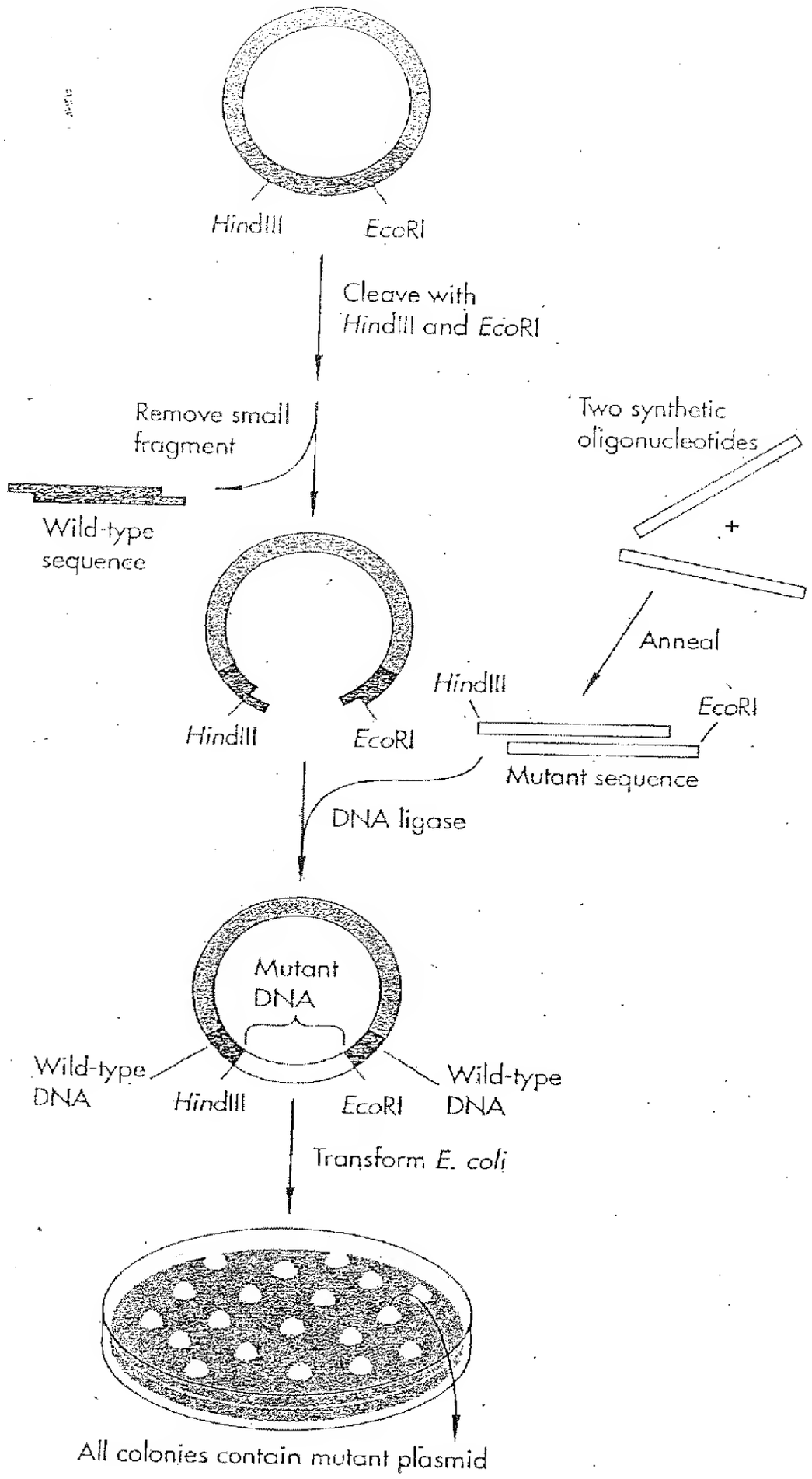


FIGURE 11-11
Mutagenesis by cassette replacement. Plasmid DNA is cleaved with restriction enzymes *Eco*RI and *Hind*III, which cut at sites that flank the sequence to be mutated. The small cleaved DNA fragment containing a portion of the wild-type sequence is removed, and a DNA fragment (cassette) containing the desired mutation is ligated into the plasmid. This mutant DNA fragment is composed of two complementary synthetic oligonucleotides that have *Eco*RI and *Hind*III sticky ends when annealed. Because there is no heteroduplex intermediate—the mutant cassette is simply swapped for the wild-type fragment—the recombinant plasmids are all mutants. A mutant cassette can be composed of degenerate oligonucleotides (see Chapter 7), resulting in a library of mutant plasmids containing different sequences.

used to study the structure of the glucocorticoid response element (GRE), an enhancer sequence that activates a family of genes in response to certain steroid hormones. The element had been mapped by deletion mutagenesis to a 30-bp region in a glucocorticoid-regulated gene. To define the sequence required for GRE function precisely, single point mutations throughout the 30-bp region were generated and tested in cells for inducibility by glucocorticoid hormone. Two complementary oligonucleotides were synthesized that carried the 30-bp GRE, but synthesis was performed under conditions in which incorrect nucleotides were incorporated at a low frequency (Figure 11-13). These "doped" oligonucleotides (that is, oligonucleotides produced by doping; see Figure 11-13) were annealed and inserted as a cassette into a promoter that lacked a GRE. Using this method, most single-nucleotide substitutions at the 30 positions were obtained. Such a collection of mutants would have been unthinkable before oligonucleotides revolutionized in vitro mutagenesis.

Gene Synthesis Facilitates Production of Normal and Mutant Proteins

The oligonucleotide-directed mutagenesis methods we have described use a single oligonucleotide or a pair of complementary oligonucleotides to insert mutant sequences into an otherwise natural DNA fragment. With the increasing availability of longer oligonucleotides, it is now feasible to assemble an entire gene from synthetic units. This is done by synthesizing a set of oligonucleotides, typically 40 to 80 nucleotides in length, that can be annealed and ligated in vitro to assemble an entire double-stranded DNA

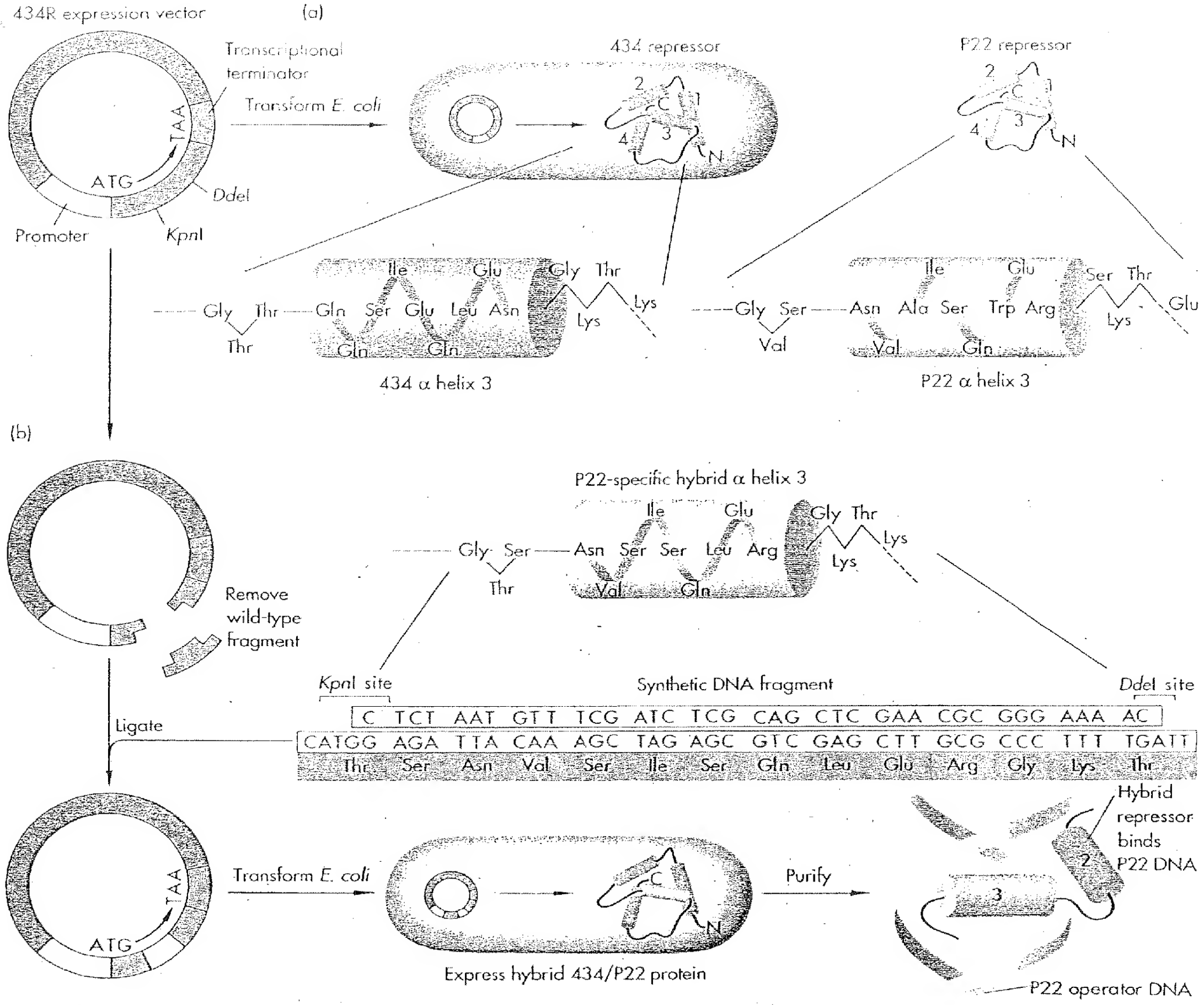


FIGURE 11-12
The helix swap experiment. Amino acids in the phage 434 repressor protein believed responsible for recognition of the 434 operator were changed by cassette mutagenesis (Figure 11-11) of 434 DNA to the amino acids believed to perform the same function in an analogous region of phage P22 repressor protein. (a) Expression in *E. coli* of the 434 repressor protein (left), with an enlargement of the site believed to bind the 434 operator; (right) the corresponding section of the P22 repressor protein. (b) A cassette was synthesized resembling the 434 domain, but with P22-type substitutions at positions thought to be essential for recognizing P22 operator DNA. This was ligated into the digested 434 plasmid, and the recombinant vector was introduced into *E. coli* to produce the hybrid protein, which then recognized P22 operator DNA but not the 434 operator.

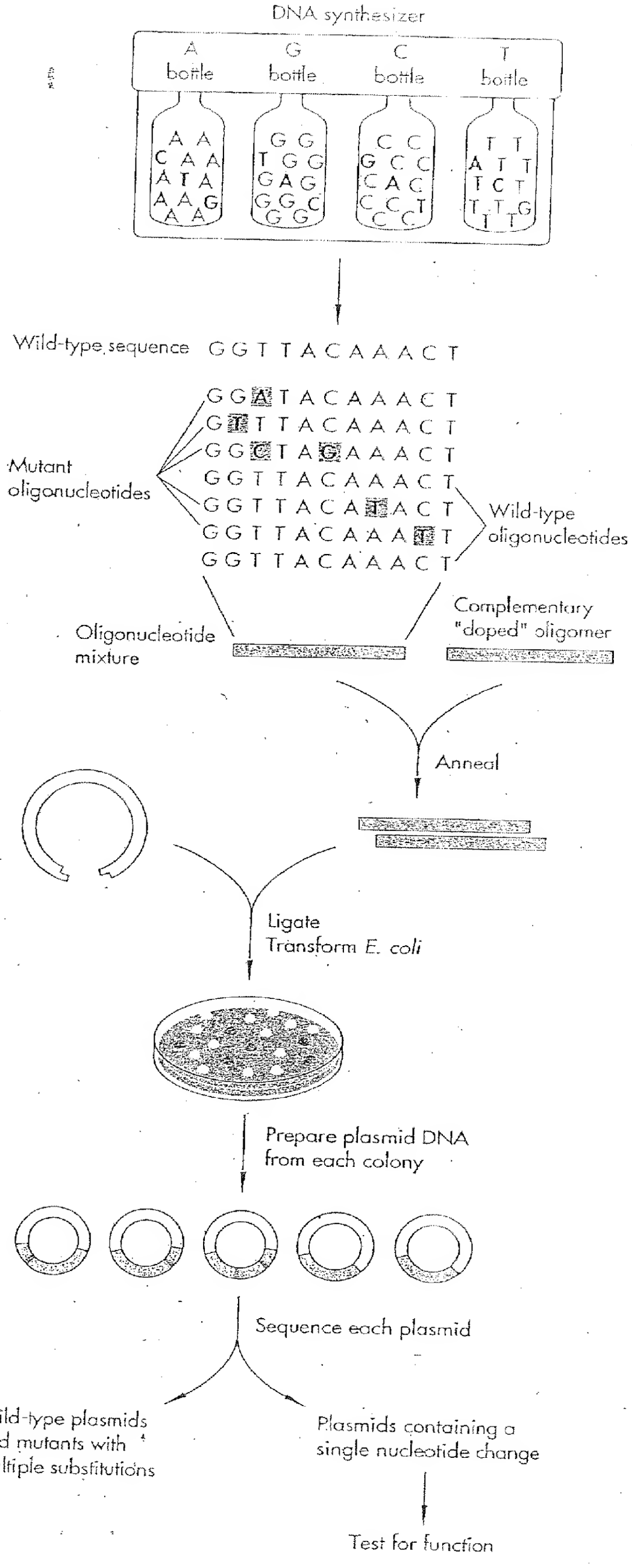


FIGURE 11-13
Cassette mutagenesis using doped oligonucleotides to generate numerous mutants in a single experiment. An oligonucleotide cassette encoding the glucocorticoid response element (GRE) was synthesized by a DNA synthesis machine. Synthesis was done under conditions in which each bottle containing a particular nucleotide precursor was "contaminated" (doped) with small amounts of the other three precursors. In the example above, the DNA synthesizer was instructed to make an oligonucleotide with the sequence GGTTACAAACT. Thus, when a nucleotide precursor is called for—a C for example—the machine adds an aliquot of the solution from the C bottle, and a C base is coupled to the end of most of the oligonucleotide chains. However, because the C bottle contains a small amount of A, G and T, an incorrect base is sometimes added instead. Since the concentration of C is roughly 30 times that of A, G and T, an incorrect base will be added to about 1 out of 30 molecules. This results in a doped collection of oligonucleotides, which actually consists of many different sequences, some wild-type and some with substitutions. The level of contamination was adjusted to favor synthesis of oligonucleotides with only one substitution, but because substitutions occur randomly, some molecules in the collection had none and others had two or more. Cassettes were formed by annealing complementary doped oligonucleotides and ligated into a vector. Plasmid DNA was isolated from 546 individual *E. coli* transformants and analyzed by sequencing. Of these, 224 were wild-type, 218 contained one substitution (for the 30 bases, of interest, 74 of the 90 possible single substitutions were recovered), and the rest contained two or more.

molecule (Figure 11-14). In gene synthesis, the experimenter has total control over the sequence of the gene. It can be wild-type or mutant in any way required. Because most amino acids are encoded by multiple triplet codons, genes encoding wild-type proteins can be constructed using different codons. Codons can be chosen to place unique restriction sites throughout the sequence so that mutant cassettes can be easily swapped in. This was done with the bacterial rhodopsin gene. Replacing a fragment of the synthetic gene with a new synthetic fragment identified the amino acid that is linked to the photon-absorbing chromophore that initiates photosynthesis. Other fragments can be exchanged as cassettes to study other important structural features of the protein.

Codons can also be changed by gene synthesis to allow production of proteins at high levels in other organisms. Studies of the biochemistry of the Fos protein, encoded by a cellular protooncogene in animal

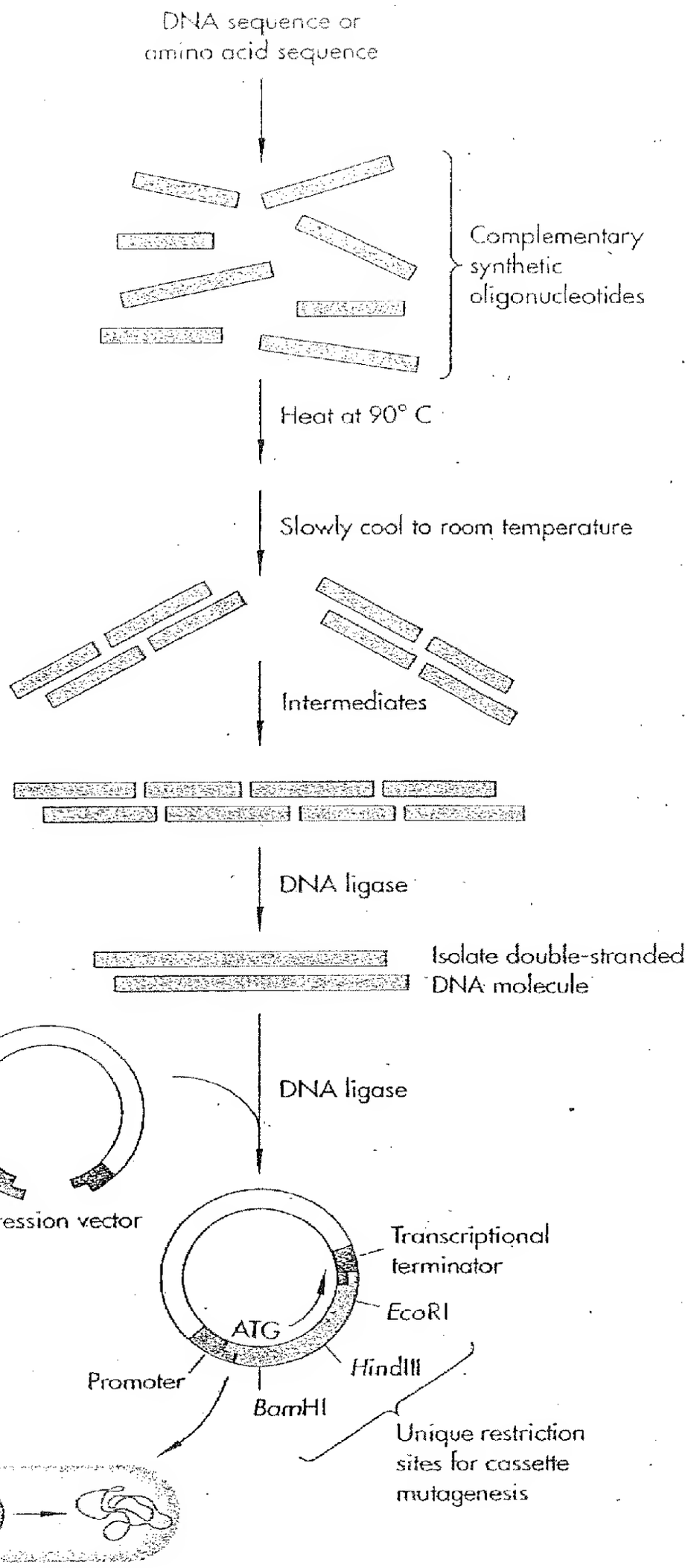


FIGURE 11-14
Gene synthesis by ligation of complementary oligonucleotides. To synthesize a gene that encodes a protein of interest, a set of overlapping complementary oligonucleorides are designed that can be combined to form a double-stranded DNA molecule that encodes the entire protein. The oligonucleotides are mixed together, heated at 90°C for a few minutes to denature the strands, and then cooled slowly to room temperature. During this period the oligonucleotides anneal through complementary base pairs. The oligonucleotides are designed so that each one anneals to two adjacent oligonucleotides from the opposite strand, bridging them. Generally, oligonucleotides ranging in length from 40 ro 80 nucleotides are used in gene synthesis. The annealed oligonucleotides are covalently linked by DNA ligase, producing two contiguous DNA strands. This synthetic gene is usually purified from a gel before ligation into a vector. The resultant recombinant plasmid is obtained following transformation into *E. coli* and is sequenced to check that the correct sequence was synthesized. The sequence of the synthetic gene can be designed to place restriction sites at convenient locations for cassette mutagenesis.

cells (Chapter 18), have been severely hampered by the inability to produce the protein in *E. coli*. This problem was finally solved by synthesizing a portion of the *fos* gene entirely from oligonucleotides, changing natural *fos* codons to the codons used most efficiently in *E. coli*. Insertion of this synthetic gene into an *E. coli* expression vector allowed for the first time the production of large quantities of active Fos protein. The gene was also designed with several unique restriction sites so that efficient cassette mutagenesis can now be coupled to the biochemical assays for Fos function.

The PCR Can Be Used to Construct Genes Encoding Chimeric Proteins

The ease with which mutations can be made in a protein coding sequence has revolutionized the study of protein function. A functional domain can be identified by making a series of mutant proteins, then testing which substitutions cause a change in function. However, it is not often easy to decide where to make a mutation. In the example of the helix swap experiment (Figure 11-12), the domain that bound DNA had been previously identified. And the design of the experiment was guided by having a model for the three-dimensional structure of the repressor protein. However, for most proteins, little structural information is available. Identifying a functional domain—for example, a region of the protein that may interact with another protein—is difficult to do by inspecting the primary amino acid sequence. A simple straregy

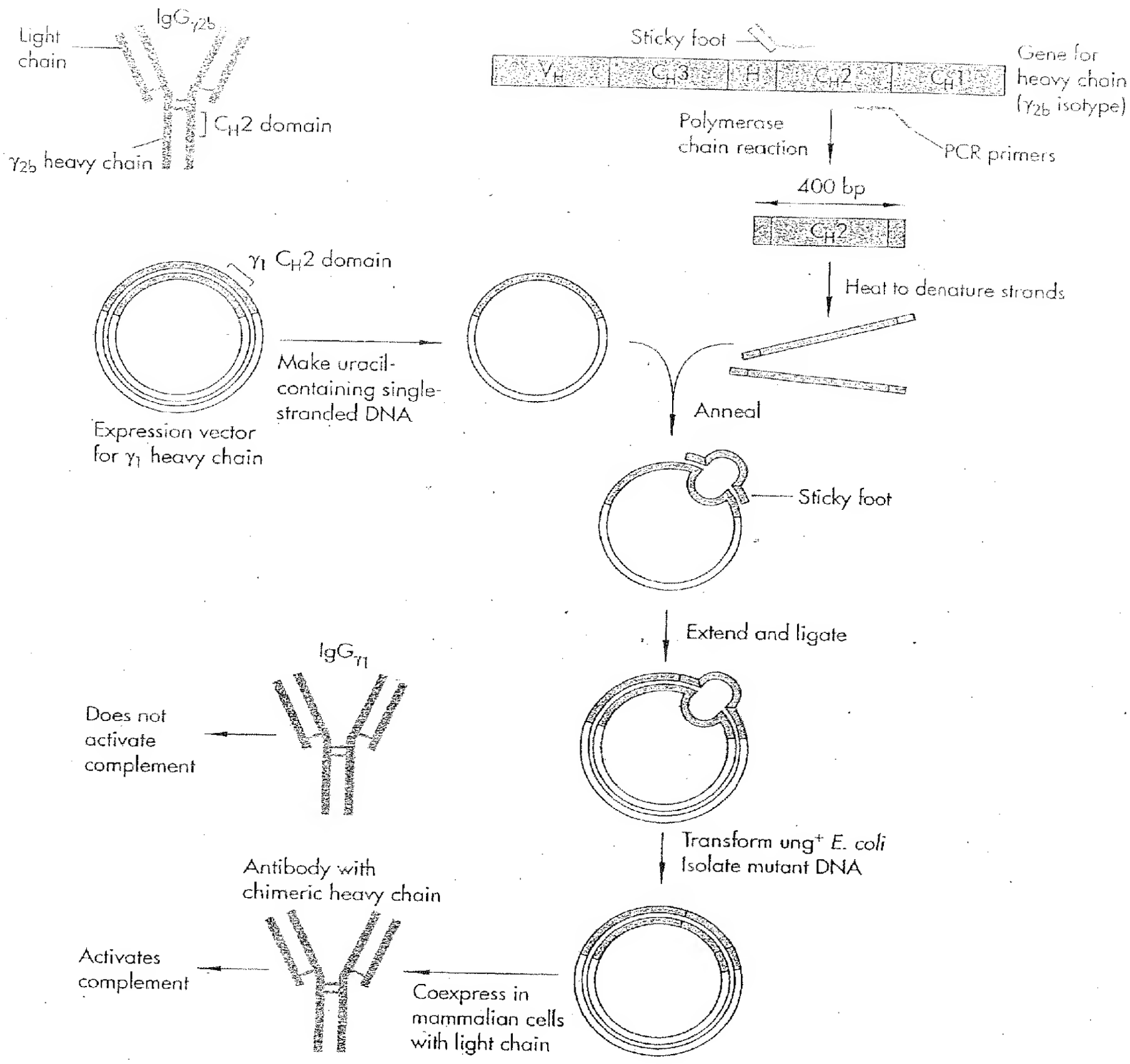


FIGURE 11-15
Construction of a chimeric antibody heavy-chain-encoding gene by "sticky feet–directed" mutagenesis. Antibodies containing a γ2b heavy chain are known to participate in complement-dependent cell lysis, whereas antibodies containing γ1 heavy chains do not. In order to identify which domain of the γ2b heavy chain is responsible for this property, an antibody containing a chimeric heavy chain was produced. To construct a gene encoding the chimeric heavy chain, a 400-bp fragment encoding the C_H2 domain from a γ1 heavy chain was replaced with the homologous segment from a γ2b gene. Since there were no convenient restriction sites at the ends of the C_H2 segments, the 400-nucleotide-long γ2b DNA was prepared by PCR. The PCR primers were complementary to the ends of the γ2b DNA but contained additional nucleotides (the sticky feet) that were complementary to γ1 DNA at the boundaries of the γ1 C_H2 domain. The strands of the PCR-generated fragment were separated by heating, then one strand was used as the primer in a mutagenesis experiment using a uracil-containing single-stranded γ1 DNA template by the method shown in Figure 11-10. The resulting chimeric heavy-chain gene was coexpressed with a light chain gene in mammalian cells to form an antibody that now activated complement. Since only the C_H2 domain came from the γ2b heavy-chain, this result demonstrated that the γ2b C_H2 domain contains the information necessary to activate complement-dependent cell lysis. Sticky feet–directed mutagenesis provided a simple means for constructing this complicated gene.

that helps to narrow down important amino acids in a protein is the analysis of chimeras between related proteins. We have previously discussed the use of computer programs to identify related proteins by comparison of their amino acid sequences (Chapter 8). Chimeric proteins are constructed by replacing a segment of one protein with the *homologous* segment from another protein. Although the two proteins have functional differences, their sequence similarity often indicates that they share a common overall structure. A striking example of this was in the analysis of human growth hormone (hGH). A series of chimeric proteins were made in which most of the amino acids were derived from hGH but which contained segments from related hormones, such as human prolactin. Using this strategy, regions of hGH that interact with the hGH receptor were identified. In Chapter 17, we will see how functional regions of a receptor which spans the membrane seven times were identified by the study of chimeras.

The 434/P22 repressor (Figure 11-12) and hGH chimeras were constructed by ligation of short oligonucleotide cassettes into the coding sequence. A different strategy (Figure 11-15) was used to prepare a chimeric antibody in which a 400-bp segment from a γ1 heavy-chain gene was replaced by the homologous segment from a γ2b gene. A 400-bp DNA fragment was generated by PCR that encoded the new sequence to be inserted and two 30-base "sticky feet" on each end. The double-stranded PCR fragment was heated to denature the two strands, and then one of the single-stranded molecules was utilized in a primer extension experiment (as in Figure 11-8). Had the gene synthesis method been employed, construction of the chimeric gene would have required twenty 40-nucleotide-long oligomers. Instead, the sticky feet method used only two oligonucleotide primers for PCR.

Mutagenesis Is the Gateway to Gene Function and Protein Engineering

It would be difficult to overestimate the importance of in vitro mutagenesis techniques to biology and biotechnology. The harnessing of enzymes that operate on DNA and the refinement of oligonucleotide synthesis have made changing gene sequences an almost trivial task. And the ability to operate on DNA lets us also change the structure of the products of genes—RNA and, most importantly, proteins. Thus, the impact of this technology is twofold. It has revolutionized how research is done in molecular biology by creating the entirely new concept of "reverse genetics"—changing gene sequence first, then examining gene function. And it opens the door to sophisticated protein engineering (see Chapter 23), the ability to make changes in natural gene products that make them do their jobs better. The impact of protein engineering on medicine and industry will be substantial.

Reading List

General

Wu, R., L. Grossman, and K. Moldave, eds. *Methods in Enzymology*, vol. 100: *Recombinant DNA, Part B.* Academic, New York, 1983.

Gait, M. *Oligonucleotide Synthesis: A Practical Approach.* IRL, Oxford, Eng., 1984.

Wu, R., and L. Grossman, eds. *Methods in Enzymology*, vol 154: *Recombinant DNA, Part E. Academic*, New York, 1987.

Original Research Papers

STRATEGIES FOR IN VITRO MUTAGENESIS

Botstein, D., and D. Shortle. "Strategies and applications of in vitro mutagenesis." *Science*, 229: 1193–1201 (1985). [Review]

Smith, M. "In vitro mutagenesis." *Ann. Rev. Gen.*, 19: 423–462 (1985). [Review]

Zoller, M. J. "New molecular biology methods for protein engineering." *Curr. Opin. Struct. Biol.*, 1: 605–610 (1991). [Review]

INSERTION MUTAGENESIS

Heffron, F., M. So, and B. J. McCarthy. "In vitro mutagenesis of a circular DNA molecule using synthetic restriction sites." *Proc. Natl. Acad. Sci. USA*, 75: 6012–6016 (1978).

Abraham, J., J. Feldman, K. A. Nasmyth, J. N. Strathern, A. J. S. Klar, J. R. Broach, and J. B. Hicks. "Sites required for position-effect regulation of mating type information in yeast." *Cold Spring Harbor Symp. Quant. Biol.*, 47: 989–998 (1983).

Rees-Jones, R. W., and S. P. Goff. "Insertional mutagenesis of the Abelson murine leukemia virus genome: identification of mutants with altered kinase activity and defective transformation ability." *J. Virol.*, 62: 978–986 (1988).

DELETION MUTAGENESIS

Lai, C. J., and D. Nathans. "Deletion mutants of SV40 generated by enzymatic excision of DNA segments from the viral genome." *J. Mol. Biol.*, 89: 179–193 (1974).

Mertz, J. E., J. Carbon, M. Herzberg, R. W. Davis, and P. Berg. "Isolation and characterization of individual clones of simian virus 40 mutants containing deletions, duplications and insertions in their DNA." *Cold Spring Harbor Symp. Quant. Biol.*, 39: 69–84 (1974).

Sakonju, S., D. Bogenhagen, and D. D. Brown. "A control region in the center of the 5S gene directs specific initiation of transcription, I: the 5′ border of the region." *Cell*, 19: 13–26 (1980).

McKnight, S. L., and R. Kingsbury. "Transcriptional control signals of a eukaryotic protein-coding gene." *Science*, 217: 316–324 (1982).

Struhl, K. "The yeast his3 promoter contains at least two distinct elements." *Proc. Natl. Acad. Sci. USA*, 79: 7385–7389 (1982).

CHEMICAL MUTAGENESIS

Chu, C. T., D. S. Parris, R. A. F. Dixon, F. E. Farber, and P. A. Schaffer. "Hydroxylamine mutagenesis of HSV DNA and DNA fragments: introduction of mutations into selected regions of the viral genome." *Virol.*, 98: 168–181 (1979).

Shortle, D., and D. Nathans. "Regulatory mutants of simian virus 40: constructed mutants with base substitutions at the origin of viral DNA replication." *J. Mol. Biol.*, 131: 801–817 (1979).

Weiher, H., and H. Schaller. "Segment-specific mutagenesis: extensive mutagenesis of a lac promoter/operator element." *Proc. Natl. Acad. Sci. USA*, 79: 1408–1412 (1982).

Myers, R. M., L. S. Lerman, and T. Maniatis. "A general method for saturation mutagenesis of cloned DNA fragments." *Science*, 229: 242–247 (1985).

ENZYMATIC MISINCORPORATION MUTAGENESIS

Shortle, D. P. Grisafi, S. J. Benkovic, and D. Botstein. "Gap misrepair mutagenesis: efficient site-directed introduction of transition, transversion, and frameshift mutations in vitro." *Proc. Natl. Acad. Sci. USA*, 79: 1588–1592 (1982).

Zakour, R. A., and L. A. Loeb. "Site-specific mutagenesis by error-directed DNA synthesis." *Nature*, 295: 708–710 (1982).

Abarzua, P., and K. J. Marians. "Enzymatic techniques for the isolation of random single-base substitutions in vitro at high frequency." *Proc. Natl. Acad. Sci. USA*, 81: 2030–2034 (1984).

Leung, D. W., E. Chen, and D. V. Goeddel. "A method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction." *Technique*, 1: 11–15 (1989)

OLIGONUCLEOTIDE-DIRECTED MUTAGENESIS

Hutchinson, C. A., Phillips, M. H. Edgell, S. Gillam, P. Jahnke, and M. Smith. "Mutagenesis at a specific position in a DNA sequence." *J. Biol. Chem.*, 253: 6551–6560 (1978).

Wallace, R. B., M. Schold, M. J. Johnson, P. Dembek, and K. Itakura. "Oligonucleotide-directed mutagenesis of the human β-globin gene: a general method for producing specific point mutations in cloned DNA." *Nuc. Acids Res.*, 9: 3642–3656 (1981).

Zoller, M. J., and M. Smith. "Oligonucleotide-directed mutagenesis using M13-derived vectors: an efficient and general procedure for production of point mutations in any fragment of DNA." *Nuc. Acids Res.*, 10: 6487–6500 (1982).

Kramer, W., V. Drutsa, H. W. Jansen, M. Pflugfelder, and H.-J. Fritz. "The gapped duplex DNA approach to oligonucleotide-directed mutation construction." *Nuc. Acids Res.*, 12: 9441–9456 (1984).

Kunkel, T. A. "Rapid and efficient site-specific mutagenesis without phenotypic selection." *Proc. Natl. Acad. Sci. USA*, 82: 477–492 (1985).

Taylor, J. W., J. Ott, and F. Eckstein. "The rapid generation of oligonucleotide-directed mutations at high frequency using phosphorothioate-modified DNA." *Nuc. Acids Res.*, 13: 8765–8785 (1985).

CASSETTE MUTAGENESIS

Lo, K.-M., S. S. Jones, N. R. Hackett, and H. G. Khorana. "Specific amino acid substitutions in bacterioopsin: replacement of a restriction fragment in the structural gene by synthetic DNA fragments containing altered codons." *Proc. Natl. Acad. Sci. USA*, 81: 2285–2289 (1984).

Wells, J. A., M. Vasser, and D. B. Powers. "Cassette mutagenesis: an efficient method for generation of multiple mutations at defined sites." *Gene*, 34: 315–323 (1985).

Wharton, R., and M. Ptashne. "Changing the binding specificity of a repressor by redesigning an alpha-helix." *Nature*, 316: 601–605 (1985).

Reidhaar-Olson, J. F., and R. T. Sauer. "Combinatorial cassette mutagenesis as a probe of the informational content of protein sequences." *Science*, 241: 53–57 (1988).

DOPED OLIGONUCLEOTIDE MUTAGENESIS

McNeil, J. B., and M. Smith. "*Saccharomyces cerevisiae* CYC1 mRNA 5'-end positioning: analysis by in vitro mutagenesis, using synthetic duplexes with random mismatch base pair." *Mol. Cell. Biol.*, 5: 3545–3551 (1985).

Hutchison, C. A., S. K. Nordeen, K. Vogt, and M. H. Edgell. "A complete library of point substitution mutations in the glucocorticoid response element of mouse mammary tumor virus." *Proc. Natl. Acad. Sci. USA*, 83: 710–714 (1986).

GENE SYNTHESIS

Ferreti, L., S. S. Karnik, H. G. Khorana, N. Nassal, and D. D. Oprian. "Total synthesis of a gene for bovine rhodopsin." *Proc. Natl. Acad. Sci. USA*, 83: 599–603 (1986).

Caruthers, M. H., A. D. Barone, S. L. Beaucage, D. R. Dodds, E. F. Fisher, L. J. McBride, M. Matteucci, Z. Stabinsky, and J.-Y. Tang. "Chemical synthesis of deoxyoligonucleotides by the phosphoramidite method." *Meth. Enzymol.*, 154: 287–313 (1987).

Abate, C., D. Luk, R. Getz, F. J. Rauscher, III, and T. Curran. "Expression and purification of the leucine zipper and DNA-binding domains of Fos and Jun: both Fos and Jun contact DNA directly." *Proc. Natl. Acad. Sci. USA*, 87: 1032–1036 (1990).

CONSTRUCTING CHIMERIC GENES BY PCR

Higuchi, R., B. Krummel, and R. K. Saiki. "A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions." *Nuc. Acids Res.*, 16: 7351–7367 (1988).

Clackson, T., and G. Winter. "'Sticky feet'-directed mutagenesis and its application to swapping antibody domains." *Nuc. Acids Res.*, 17: 10163–10170 (1989).

PROTEIN STRUCTURE AND FUNCTION

Winter, G., A. R. Fersht, A. J. Wilkinson, M. Zoller, and M. Smith. "Redesigning enzyme structure by site-directed mutagenesis." *Nature*, 299: 756–758 (1982).

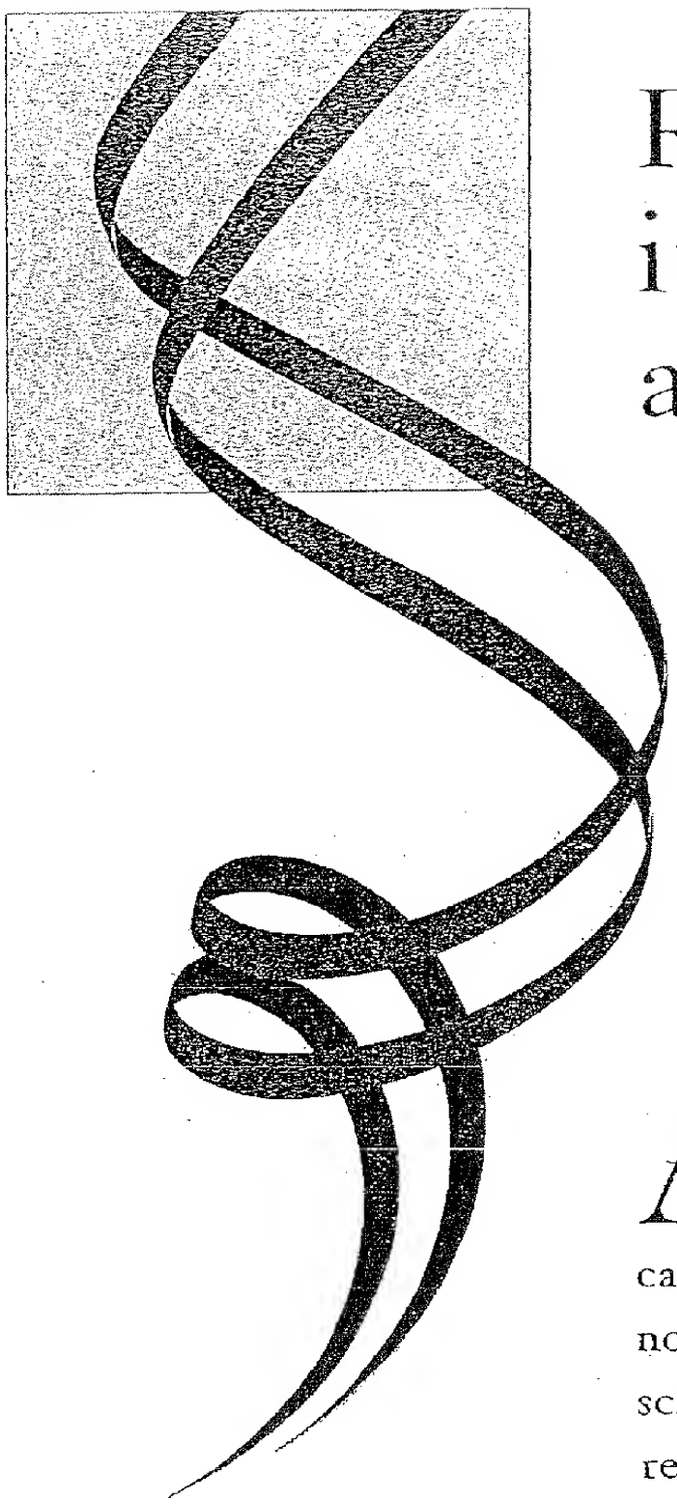
Cunningham, B. C., P. Jhurani, P. Ng, and J. A. Wells. "Receptor and antibody epitopes in human growth hormone identified by homolog-scanning mutagenesis." *Science*, 243: 1330–1336 (1989).

Gibbs, C. S., and M. J. Zoller. "Rational scanning mutagenesis of a protein kinase identifies functional regions involved in catalytic and substrate interactions." *J. Biol. Chem.*, 266: 8923–8931 (1991).

CHAPTER

23

Recombinant DNA in Medicine and Industry



As soon as the first successful cloning experiments were reported in 1973, applications for this powerful technology quickly followed. The significance of being able to produce large quantities of human proteins that were normally available in exceedingly small amounts, if at all, was not lost on scientists, physicians, and businessmen alike. In 1976 *biotechnology* became a reality as the methodologies for DNA cloning, oligonucleotide synthesis, and gene expression converged in a single experiment, in which a human protein was expressed from recombinant DNA for the first time. The protein was somatostatin, a 14 amino acid peptide neurotransmitter. The gene encoding somatostatin was not the natural gene but was synthesized chemically and cloned into a plasmid vector for expression in *E. coli*. Soon after followed the successful expression of human insulin for the treatment of diabetes, the first commercial product of the biotechnology industry. Instead of insulin extracted from the pancreases of pigs and cows, diabetics could now receive insulin identical to that normally produced by humans.

The ability to achieve such feats relied on the successes in all areas of molecular biology, including oligonucleotide synthesis, isolation of enzymes that cleave and join DNA, characterization of bacterial plasmids, and an understanding of gene expression. These methods have, of course, revolutionized research in

biology and medicine, but what is equally important, they have spawned an entirely new industry, one devoted to the cloning and production of proteins of importance to both medicine and industry. Today, proteins are produced through recombinant DNA technology for treatment of numerous diseases—cancer, allergies, autoimmune disease, neurological disorders, heart attacks, blood disorders, infections, wounds, and genetic diseases—as well as for more prosaic tasks, such as use in laundry detergents and food production. In addition, entirely new approaches to drug design have emerged from recombinant DNA technology, as scientists have gained the ability to tinker with natural proteins to improve their function and to change them in subtle and useful ways.

Expression Systems Are Developed to Produce Recombinant Proteins

Cloning the gene or cDNA encoding a particular protein is only the first of many steps needed to produce a recombinant protein for medical or industrial use. The next step is to put the gene into a host cell for production. The development of expression systems has been an important research area in both industrial and academic laboratories. The most popular expression systems are the bacteria *E. coli* and *Bacillus subtilis*, yeast, and cultured insect and mammalian cells. We have learned in earlier chapters about the development of vectors and DNA transformation methods for these organisms. Here we will discuss the issues that are important for protein production. The choice of which cell is used depends on the project goals and on the properties of the protein to be produced.

Bacterial cells offer simplicity, short generation times, and large yields of product with low costs. And, especially with *B. subtilis*, the cells can be induced to secrete the product into the culture medium, thus greatly simplifying the task of purification. But expression in prokaryotic cells has several drawbacks. Although some proteins are expressed to high levels (greater than 10 percent of the mass of all bacterial proteins), they often fail to fold properly and hence form insoluble *inclusion bodies*. Protein extracted from these inclusion bodies is often biologically inactive. Small proteins can sometimes be refolded into their active forms, but larger proteins usually cannot. A second problem is that foreign proteins are sometimes toxic to bacteria, so cell cultures producing the protein cannot be grown to high densities. This problem can often be circumvented by using an inducible promoter that is turned on to begin transcription of the gene for the foreign protein only after the culture has been grown. Third, bacterial cells lack enzymes that are present in eukaryotic cells and add posttranslational modifications, such as phosphates and sugars, to proteins. These modifications are often required for proper functioning of proteins. Researchers are addressing this problem by purifying the eukaryotic enzymes that carry out these modifications and using these enzymes to add the needed modifications to bacterially expressed proteins.

Yeast has been used for centuries by brewers and bakers, and now it toils for biotechnologists as well. As discussed in Chapter 13, yeast is a simple eukaryote that resembles mammalian cells in many ways but can be grown as quickly and cheaply as bacteria can. Yeast perform many of the posttranslational modifications found on human proteins and can be induced to secrete certain proteins into the growth medium for harvesting. A disadvantage of yeast is the presence of active proteases that degrade foreign proteins, thereby reducing the yield of product. Researchers are dealing with this problem, however, by constructing yeast strains in which the protease genes have been deleted.

Expression of heterologous proteins in insect cells by baculovirus vectors (as previously described in Figure 12-12) is a relatively new approach. The main advantages are high-level expression, correct folding, and posttranslational modifications similar to those in mammalian cells. A vaccine for the AIDS virus has been prepared by producing one of the HIV glycoproteins with this system. Although the cost of culturing insect cells is currently more than that for culturing bacteria and yeast, it is less than that for culturing mammalian cells.

Despite the significant advantages of producing human proteins in heterologous host cells, in some cases the best place to produce a mammalian protein is in mammalian cells. Great improvements have been made to promoters, vectors, transformation protocols, and host cell systems. Transient expression in mam-

malian cells (described in Figure 12-4) is often used for checking the function of a newly cloned gene and as a quick method for assessing the function of engineered proteins. The extracellular domains of cell-surface receptors (Chapter 17) have been engineered for secretion from cells by introducing a stop codon into the gene before the transmembrane domain sequence. These *soluble receptors* are valuable reagents for studying ligand binding in vitro and for screening for receptor agonists or antagonists, and they may eventually be used as therapeutics themselves. Although transient systems yield enough protein for laboratory experiments, stably integrated amplified genes in mammalian cells are used for the large-scale production of proteins such as tissue plasminogen activator, which we describe later.

Insulin Is the First Recombinant Drug Licensed for Human Use

The first licensed drug produced through genetic engineering was human insulin. An important hormone that regulates sugar metabolism, insulin is produced by a small number of cells in the pancreas and secreted into the bloodstream. An inability to produce insulin results in diabetes, but daily injections of insulin are sufficient to reverse or at least allay the debilitating effects of the disease. Prior to production of the recombinant molecule, insulin for treatment of diabetes was obtained from the pancreases of pigs and cows. Although this insulin is biologically active in humans, the amino acid sequences are not identical to that of the human molecule. Thus, some patients produced antibodies against injected insulin, occasionally resulting in serious immune reactions. Because recombinant human insulin is identical to the natural product, immunogenicity should not be a problem.

In mammals, insulin is expressed as a single-chain *prepro-hormone,* which is secreted through the plasma membrane. A prepro-hormone contains extra amino acids not present in the mature hormone. Amino-terminal amino acids form the *pre* sequence and target the expressed protein for secretion. The *pro* sequence is a stretch of amino acids in the middle of the hormone sequence that is important for folding the polypeptide chain into the correct structure. During secretion, these extra amino acids are cleaved from the prepro-hormone by cellular proteases to release the mature insulin molecule, consisting of two short polypeptide chains, A and B, linked by two disulfide bonds. The principal challenge in the production of recombinant insulin was getting insulin assembled into this mature form. The initial approach was to construct synthetic genes from oligonucleotides that separately encoded the A and B chains. These were individually inserted into the *E. coli* gene encoding β-galactosidase, so the bacteria produced large fusion proteins that had the insulin sequences tacked onto the end of the β-galactosidase enzyme (Figure 23-1). These large proteins were purified from bacterial extracts, and the insulin chains were released by treatment with cyanogen bromide, a chemical that cleaves peptide bonds following methionine residues. Because a methionine codon had been inserted at the boundaries between β-galactosidase and the insulin chains in the fusion proteins, cyanogen bromide treatment clipped intact insulin chains off the fusion proteins. These were purified, mixed, and reconstituted into an active insulin molecule. This approach was refined by producing a single β-galactosidase–insulin fusion protein, which could be cleaved in a single step to release mature insulin. A similar method is now in use for the commercial production of recombinant insulin.

Recombinant Human Growth Hormone is Produced in Bacteria by Two Methods

Growth hormone is a 191 amino acid protein that is produced in the pituitary gland and regulates growth and development. Children born with growth hormone deficiency—hypopituitary dwarfs—never achieve normal stature. Regular injections of growth hormone stimulate the growth of these children so that they reach near-normal heights. Unlike the situation with insulin, animal-derived growth hormones are ineffective. Only the human protein works, and for many years it was painstakingly extracted from the pituitaries of human cadavers. One unforeseen and

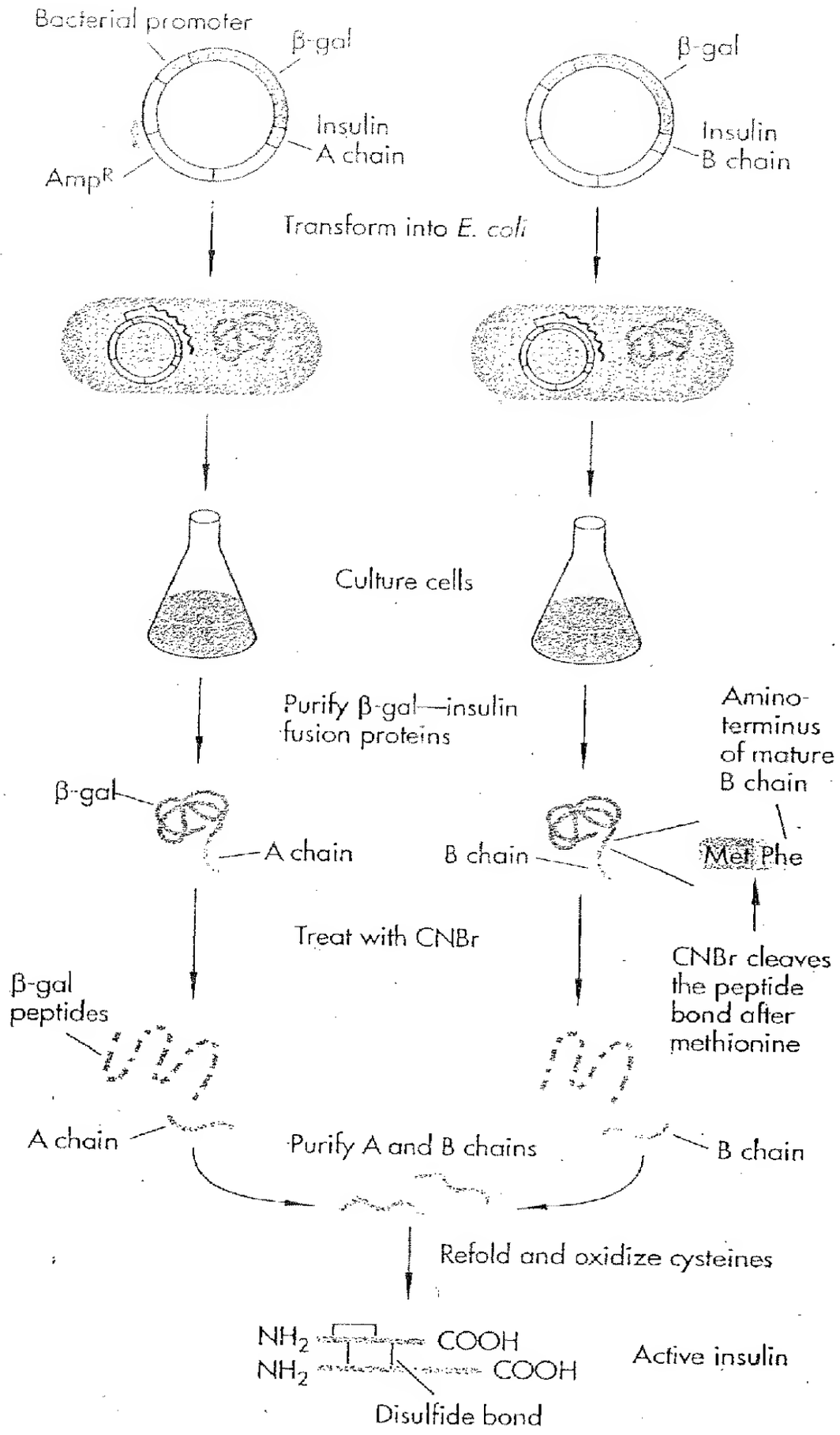


FIGURE 23-1
Expression of human insulin in *E. coli*. Recombinant insulin was first made by expressing the A and B chains separately, then refolding them into a mature insulin molecule. A DNA fragment encoding each insulin chain was made by annealing two complementary oligonucleotides that had been chemically synthesized. Each fragment was ligated into a bacterial expression vector so that, when translated, the insulin chain would be fused to the carboxy terminus of the enzyme β-galactosidase (β-gal). The expression vectors were transformed into *E. coli*, and the β-gal–insulin fusion proteins accumulated inside the bacterial cells. The cells were harvested, and each β-gal–insulin fusion protein was purified. The insulin-coding DNA was synthesized so that it started with a methionine codon. This setup provided a way to cleave off the β-gal part from the insulin polypeptide. Treatment of the fusion protein with the chemical cyanogen bromide (CNBr) results in cleavage of peptide bonds after all methionines. In this way, the natural insulin peptides were obtained. Because β-gal contains other methionine residues, CNBr treatment cleaved it into many small peptides. The insulin chains were not cleaved further because they did not contain internal methionines. The A and B chains were purified and then mixed together to form active recombinant insulin.

unfortunate consequence of growth hormone treatment, however, was the infection of a number of children with a fatal virus from one of the cadavers. Production of recombinant human growth hormone (hGH) would clearly provide a safe, reliable, and plentiful source of this drug.

The initial production of hGH was achieved by constructing a hybrid gene from the natural hGH cDNA and synthetic oligonucleotides that encoded the amino terminus of the mature form of the protein (Figure 23-2a). This coding sequence was ligated into a plasmid adjacent to a bacterial promoter. Like insulin, hGH is normally produced as a larger precursor protein containing an amino-terminal signal sequence. Because the human signal sequence would not be recognized by the bacterial secretion machinery, the 5′ end of the cDNA was reengineered with a synthetic DNA sequence enabling the bacteria to produce a nearly normal version of the mature human protein.

The first hGH expression vectors directed the production of the protein inside the cell. Purification required many steps to separate hGH from the thousands of intracellular bacterial proteins. Another way to produce the protein in bacteria is to engineer the protein so it is secreted. This can be done by linking the coding sequence for the desired protein to a signal sequence from a secreted bacterial protein, thus forming a *prehormone* (Figure 23-2b). Human growth hormone is produced by the bacteria and then secreted with the concomitant removal of the signal peptide by a bacterial protease. Secretion into the periplasm, where there are fewer proteins than inside the cell, makes purification simpler. The only difference between the secreted hGH and that produced intracellularly is the presence of an amino-terminal methionine on the intracellularly expressed molecule. Because the secreted

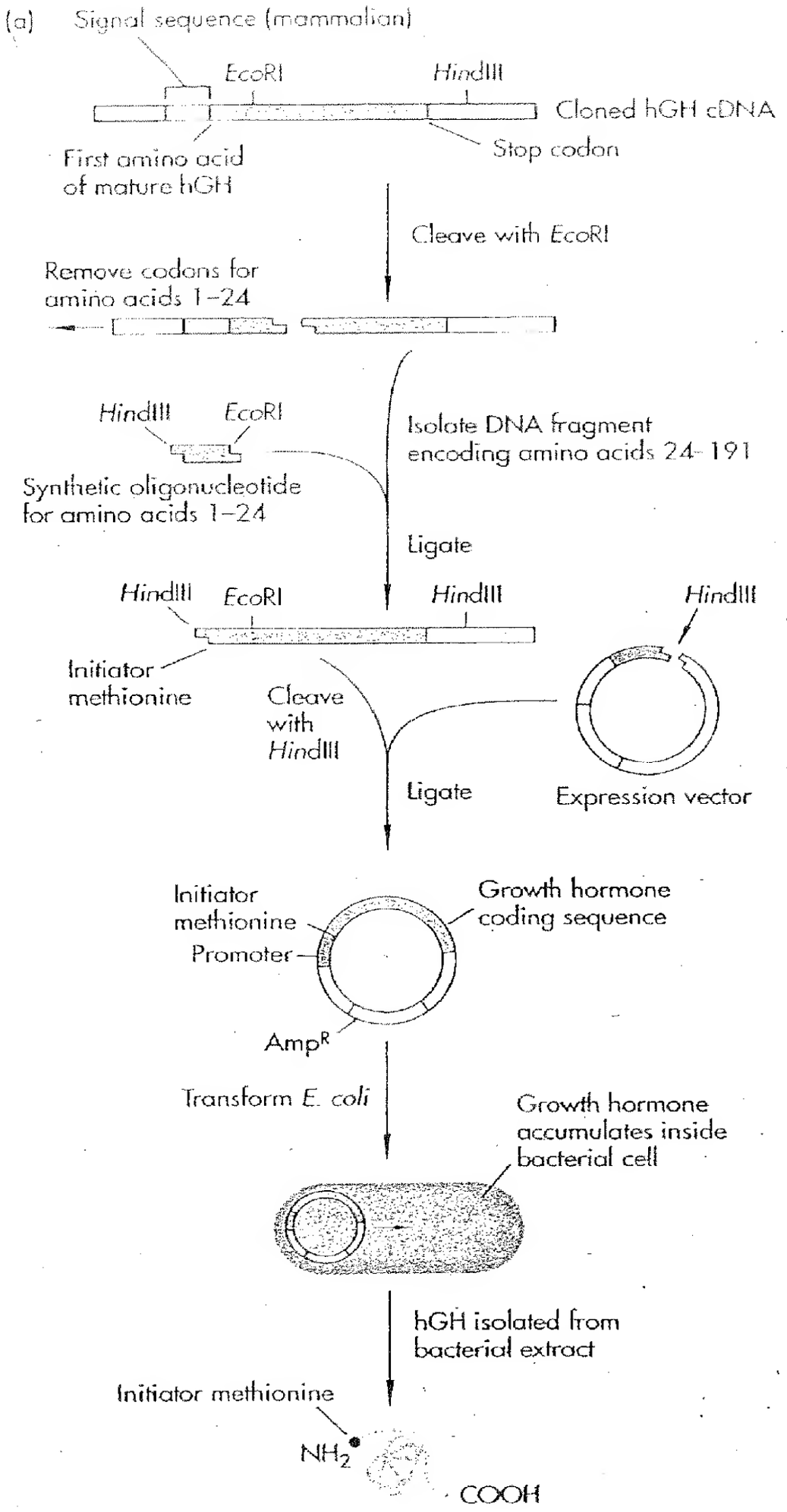


FIGURE 23-2
Bacterial production of human growth hormone (hGH). (a) An expression vector was constructed for intracellular production of hGH. The coding sequence was constructed by isolating from the cDNA a DNA fragment that encoded amino acids 24–191 and ligating this to a synthetic oligonucleotide fragment that encoded amino acids 1–24. Following introduction of the expression vector into bacterial cells, recombinant hGH was produced inside the cells. The expressed protein behaved just like natural human growth hormone but contained the initiator methionine at the amino terminus. (b) A protein can be produced in bacteria without this extra methionine by targeting it for secretion. To do this, a DNA fragment encoding a bacterial *signal sequence,* which specifies secretion of a bacterial protein, was placed in front of the hGH coding sequence. Upon introduction of this vector into bacteria, hGH is produced, and the signal sequence targets the protein for secretion. The protein accumulates in the *periplasmic space* between the inner and outer bacterial membranes and can be released by hypotonic disruption of the outer membrane. In contrast to the intracellular form of hGH, the protein produced by this procedure does not contain an initiator methionine, since a periplasmic protease cleaved off the signal sequence.

form lacks this methionine, it is called *met-less* hGH. Bacterially expressed hGH has been administered to thousands of growth hormone–deficient children, who have benefited greatly from this recombinant drug.

A Hepatitis B Virus Vaccine Is Produced in Yeast by Expression of a Viral Surface Antigen

One of the successes of modern medicine is the development and implementation of vaccines against infectious diseases. Prior to the advent of recombinant DNA technology, two types of vaccines were used. *Inactivated* vaccines are chemically killed derivatives of the actual infectious agent. *Attenuated* vaccines are live viruses or bacteria altered so that they no longer multiply in the inoculated organism. Both types of vaccines work by presenting surface proteins (antigens) to B and T lymphocytes, which become primed to respond rapidly should the organism actually become infected, usually destroying the infectious agent before any damage is done (Chapter 16). However, these types of vaccines are potentially dangerous because they can be contaminated with infectious organisms. For example, a small number of children each year contract polio from their polio vaccinations. Thus, one of the most promising applications of recombinant DNA technology is the production of *subunit vaccines*, consisting solely of the surface protein to which the immune system responds. With a subunit vaccine, there is no risk of infection.

The first successful subunit vaccine was produced for hepatitis B virus (HBV), which infects the liver and causes liver damage and, in some cases, cancer. The virus particle is coated with a surface antigen, HBsAg, and infected patients carry large aggregates of this protein in their blood. Early experiments suggested that these aggregates would make a potent vaccine, but how could they be produced in quantities sufficient to vaccinate large populations against HBV? With the cloning of the HBV genome, the possibility of a subunit vaccine could be explored. Initial attempts to produce the HBsAg protein in *E. coli* failed, so researchers turned to yeast. The HBsAg gene was inserted into a high-copy yeast expression vector (Figure 13-3) and engineered, in this case, so that it would not be secreted (Figure 23-3). Yeast transformed with this plasmid produced large quantities of the viral protein (about 1–2 percent of the total yeast protein). By growing the yeast in large fermentors, it was possible to produce 50–100 mg of the protein per liter of culture. This recombinant protein closely resembled the natural viral protein; it even formed aggregates with properties similar to those of the immunogenic aggregates found in infected patients. The yeast protein is now used commercially to vaccinate people against HBV infection.

Vaccines against many human and animal pathogens are currently in various stages of development. Recombinant DNA technology has provided a safe means to work with and to inoculate children and adults with only noninfectious parts of infectious agents. In Chapter 25, we will discuss various strategies for the development of a vaccine against the AIDS virus.

Complex Human Proteins Are Produced by Large-Scale Mammalian Cell Culture

Most of the recombinant proteins we have discussed thus far in this chapter are relatively small and simple in both structure and function. Other proteins of medical interest are considerably more complicated in structure and function, and biologically active proteins have proved difficult to produce in bacteria and yeast. In these cases, biotechnology companies have resorted to using mammalian cells for protein production. Mammalian cells are finicky and expensive to grow, but they can be counted on to produce correctly modified, fully active proteins. Thus, much effort in the biotechnology industry has been devoted to setting up fermentor systems for large-scale culture of mammalian cells.

The first drug to be produced commercially by mammalian cell culture was *tissue plasminogen activator* or *tPA*, which is administered to heart attack victims. Tissue plasminogen activator is a protease, an enzyme that cleaves other proteins. It works by clipping *plasminogen*, an inactive precursor protein, to form *plasmin*, itself a potent protease that degrades *fibrin*, the protein

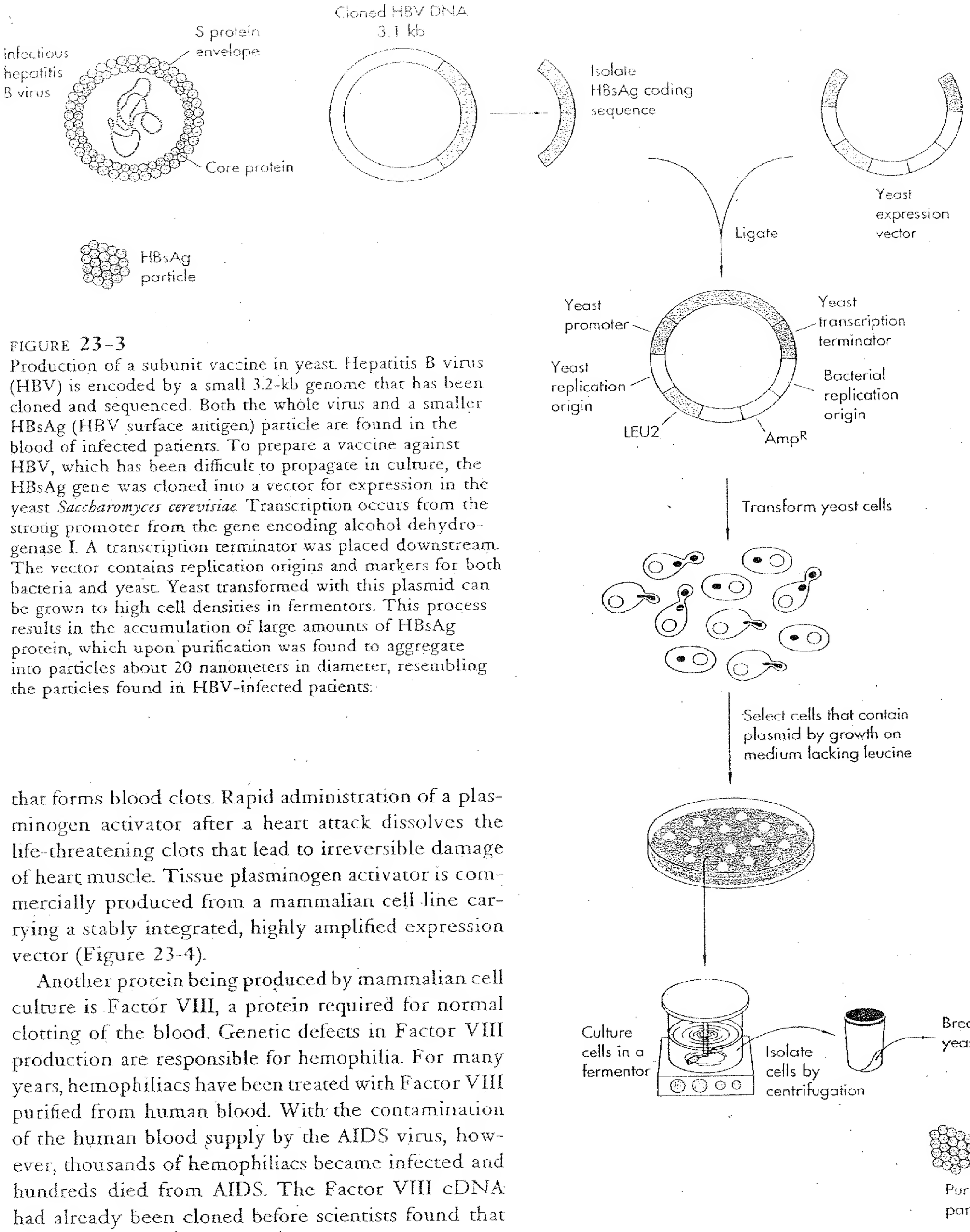


FIGURE 23–3
Production of a subunit vaccine in yeast. Hepatitis B virus (HBV) is encoded by a small 3.2-kb genome that has been cloned and sequenced. Both the whole virus and a smaller HBsAg (HBV surface antigen) particle are found in the blood of infected patients. To prepare a vaccine against HBV, which has been difficult to propagate in culture, the HBsAg gene was cloned into a vector for expression in the yeast *Saccharomyces cerevisiae*. Transcription occurs from the strong promoter from the gene encoding alcohol dehydrogenase I. A transcription terminator was placed downstream. The vector contains replication origins and markers for both bacteria and yeast. Yeast transformed with this plasmid can be grown to high cell densities in fermentors. This process results in the accumulation of large amounts of HBsAg protein, which upon purification was found to aggregate into particles about 20 nanometers in diameter, resembling the particles found in HBV-infected patients.

that forms blood clots. Rapid administration of a plasminogen activator after a heart attack dissolves the life-threatening clots that lead to irreversible damage of heart muscle. Tissue plasminogen activator is commercially produced from a mammalian cell line carrying a stably integrated, highly amplified expression vector (Figure 23–4).

Another protein being produced by mammalian cell culture is Factor VIII, a protein required for normal clotting of the blood. Genetic defects in Factor VIII production are responsible for hemophilia. For many years, hemophiliacs have been treated with Factor VIII purified from human blood. With the contamination of the human blood supply by the AIDS virus, however, thousands of hemophiliacs became infected and hundreds died from AIDS. The Factor VIII cDNA had already been cloned before scientists found that

the blood supply was contaminated with the AIDS virus. Recognition of the need for a safer source of Factor VIII accelerated efforts already under way to produce the protein using recombinant DNA methods. Like tPA, Factor VIII is a large and complex protein and can only be efficiently produced in mammalian cell culture. But the availability of recombinant protein will spare future generations of hemophiliacs from infectious agents that contaminate the blood supply.

Monoclonal Antibodies Function as "Magic Bullets"

We have discussed the use of biotechnology to produce novel vaccines that elicit antibody production by the body's immune system. As we learned in Chapter 16, antibodies are exquisitely selective proteins that can bind to a single target among millions of irrelevant sites. Researchers have long dreamed of harnessing the specificity of antibodies for a variety of uses that require the targeting of drugs and other treatments to particular sites in the body. It is this use of antibodies as targeting devices that led to the concept of the "magic bullet," a treatment that could effectively seek and destroy tumor cells and infectious agents wherever they resided.

The major limitation in the therapeutic use of antibodies is producing a useful antibody in large quantities. Initially, researchers screened *myelomas,* which are antibody-secreting tumors, for the production of

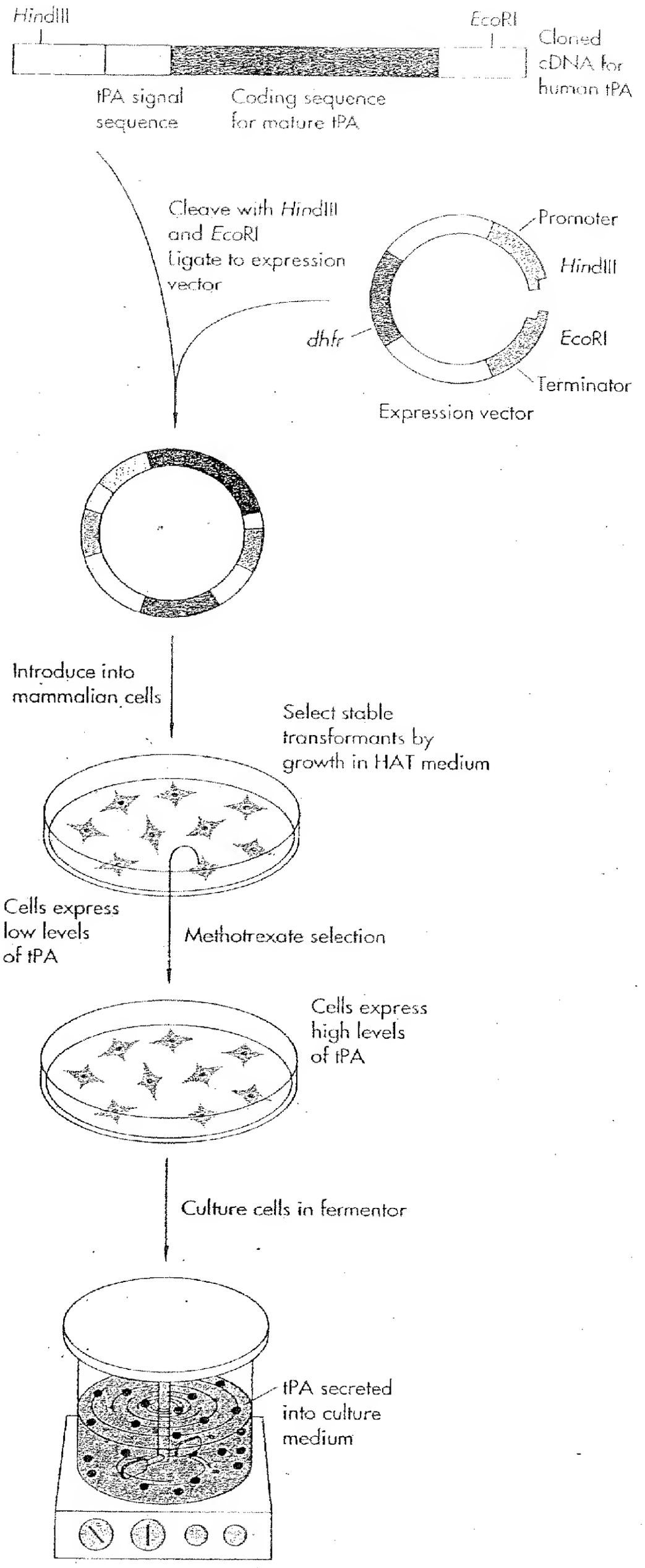


FIGURE 23-4
Production of tissue plasminogen activator (tPA) by mammalian cell culture. The cloned cDNA for human tPA was ligated into an expression vector that contained a strong promoter and terminator. The vector was stably transfected into a mammalian cell line. The initial transformants secreted tPA into the culture medium, but the level of expression was very low. Cell lines that expressed tPA to high levels were obtained using methotrexate treatment, which selects for cells that have amplified the *dhfr* gene resident in the vector together with the linked tPA expression cassette (Chapter 12). High-expressing lines are grown in large fermentors and recombinant tPA is purified from the culture medium.

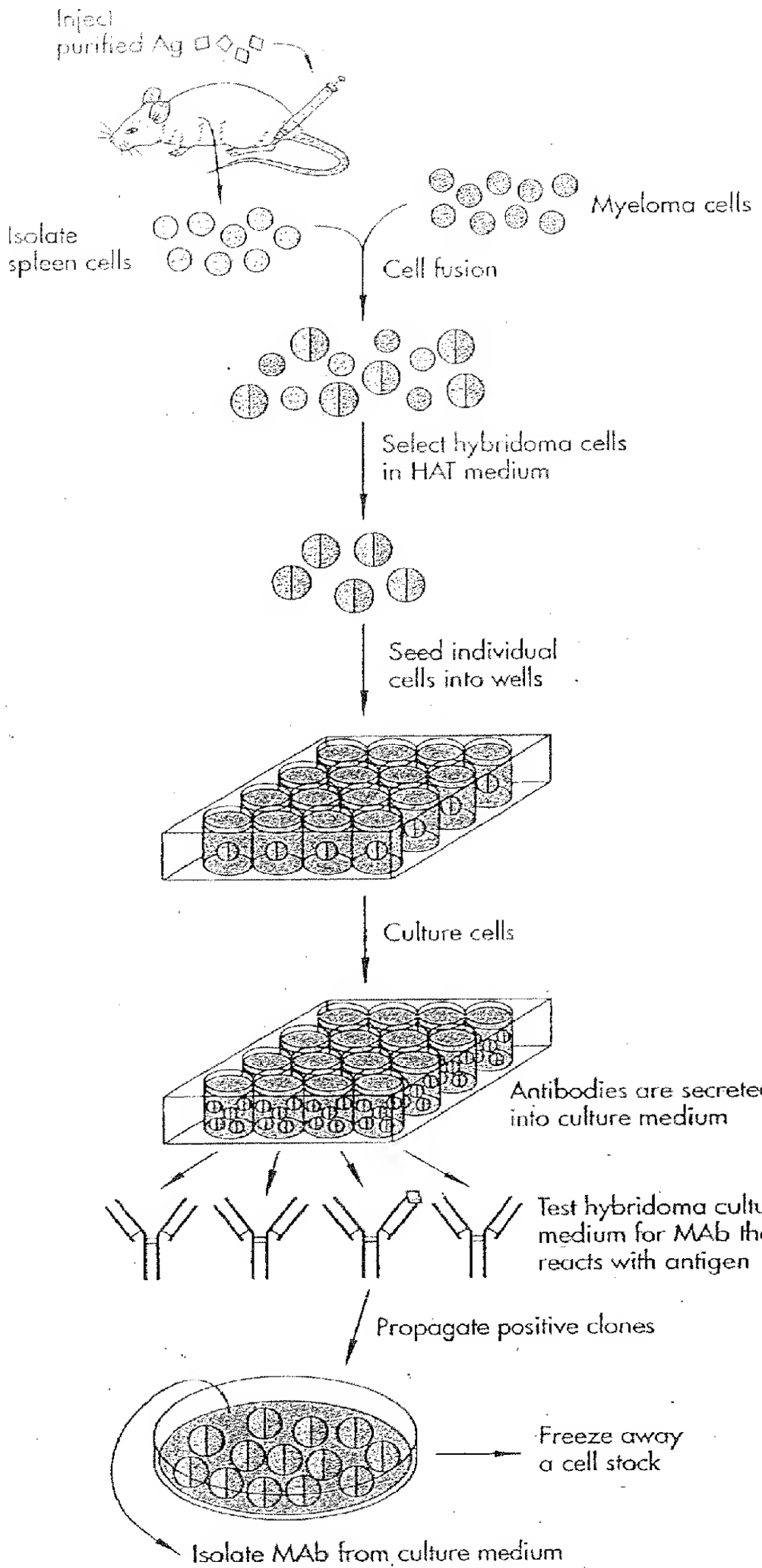


FIGURE 23-5
Production of a monoclonal antibody (MAb). A mouse is inoculated with an antigen (Ag) of interest. This stimulates the proliferation of lymphocytes expressing antibodies against the antigen. Lymphocytes are taken from the spleen and fused to myeloma cells by treatment with polyethylene glycol. Hybrid cells are selected by growth in HAT medium (Chapter 12). The myeloma cells lack the enzyme HPRT and thus die in this medium unless they become fused with a lymphocyte, which expresses the missing enzyme. Unfused lymphocyte cells soon die off as well, because they do not grow for long in culture. Individual hybrid cells are transferred to the wells of a microtiter dish and cultured for several days. Aliquots of the culture fluids are removed and tested for the presence of antibody (Ab) that binds the antigen. Cells that test positive are cultured for monoclonal antibody production. Antibody-producing cell lines are stored frozen in liquid nitrogen (this process is called *cell banking*). Aliquots can be thawed out and cultured as needed.

useful antibodies. But they lacked a means to program a myeloma to produce an antibody to their specifications. This situation changed dramatically with the development of *monoclonal antibody* technology. The procedure for producing monoclonal antibodies, or MAbs, is shown in Figure 23-5. First, a mouse or rat is inoculated with the antigen to which an antibody is desired. After the animal mounts an immune response to the antigen, its spleen, which houses antibody-producing cells (lymphocytes), is removed, and the spleen cells are fused en masse to a specialized myeloma cell line that no longer produces an antibody of its own. The resulting fused cells, or *hybridomas*, retain properties of both parents. They grow continuously and rapidly in culture like the myeloma cell, yet they produce antibodies specified by the lymphocyte from the immunized animal. Hundreds of hybridomas can be produced from a single fusion experiment, and they are systematically screened to identify those producing large amounts of a desired antibody. Once identified, this antibody is available in limitless quantities. Monoclonal antibodies are already widely used for the diagnosis of infections and cancer and for the imaging of tumors for radiotherapy. And investigations into their use in the direct treatment of cancer, inflammation, and immune disorders is on the rise.

Human Antibodies That Recognize Specific Antigens Can Be Directly Cloned and Selected

One new application of monoclonal antibody technology is the generation of *abzymes*, antibodies that behave like enzymes to catalyze a chemical reaction.

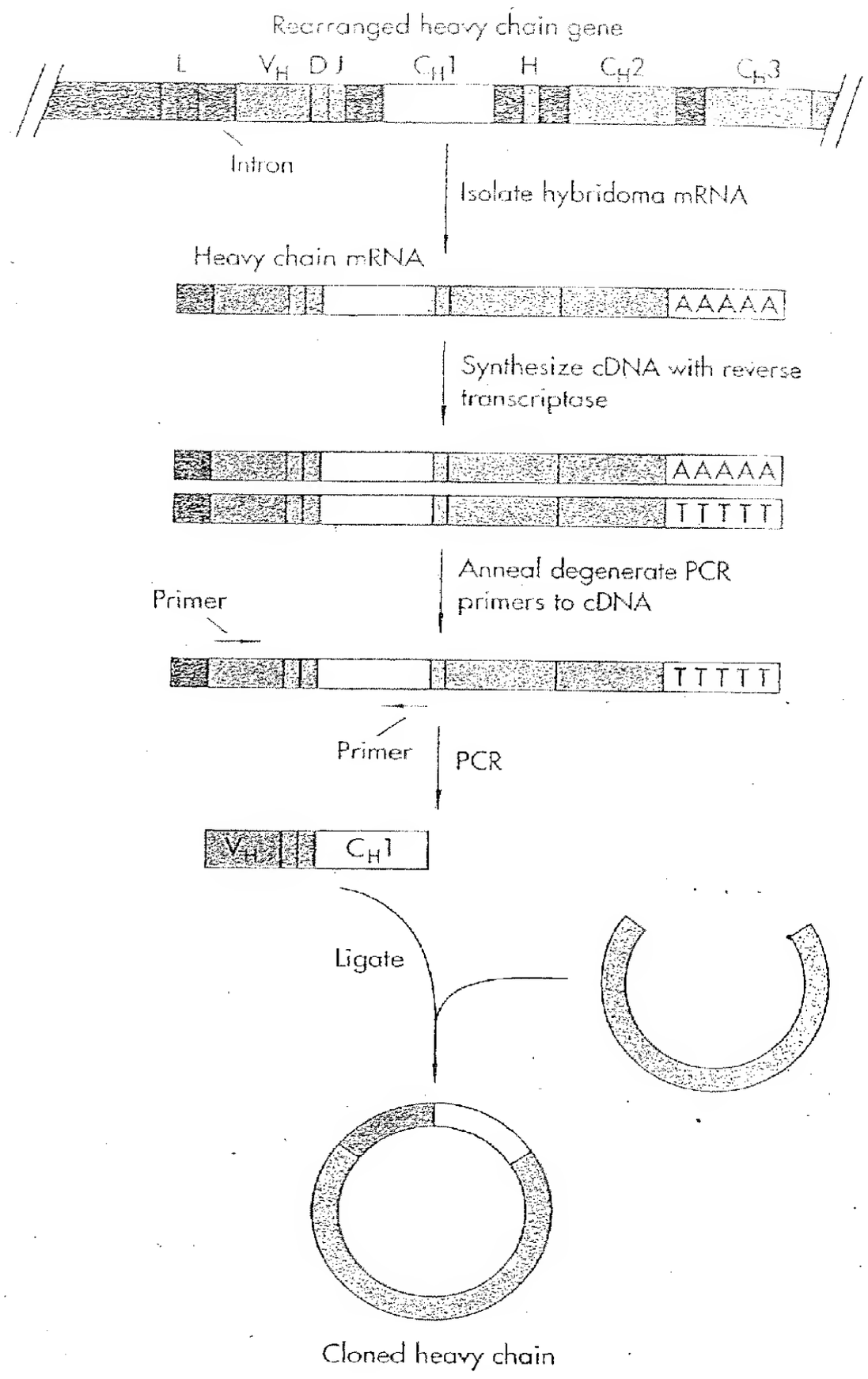


FIGURE 23-6
Direct cloning of antibody cDNAs by PCR. To engineer an antibody, the amino acid sequence of the variable domain needs to be determined. This could be done by sequencing a purified preparation of the heavy- (H) and light- (L) chain proteins, but a simpler method is to deduce the sequence from the cloned cDNA. In the past, a cDNA library was prepared from hybridoma mRNA and screened with probes from the constant regions of the H and L chain genes. A simpler method has been developed that uses the PCR. From a comparison of a large number of antibody sequences, amino acids frequently found at the amino termini of antibodies were identified. From this information, a set of degenerate PCR primers was designed that correspond to all the possible sequences in this region. Because the amino acids in the constant domains of different antibodies are nearly identical, only one PCR primer is needed for the 3′ end of each H and L chain sequence. To directly clone the antibody cDNAs, cDNA is prepared by treating hybridoma mRNA with reverse transcriptase, mixed with a pair of PCR primers (in this case, for amplifying the heavy chain sequences), and subjected to PCR. Without knowledge of the amino terminus of the antibody chain, a PCR had to be set up with each of the different 5′ primers until an amplified DNA fragment was obtained. The process can be simplified if the sequence of the first six or seven amino acids of the antibody can be determined; this is sufficient to design a single 5′ PCR primer.

Enzymes catalyze reactions by stabilizing a chemical structure intermediate between the substrate and product, termed the *transition state*. Thus, if monoclonal antibodies could be made to a transition state analogue—a molecule resembling the transition state of a chemical reaction—then some of these antibodies might have catalytic activity. The ability to produce custom-designed catalysts would be very valuable, especially to the chemical and pharmaceutical industries.

Initial attempts to produce catalytic antibodies indicated that they were exceedingly rare and often not found among the hybridomas produced by conventional monoclonal antibody technology. An excellent fusion might produce several hundred different antibodies, but the entire repertoire of antibodies that can be produced by the immune system is perhaps 100 million. How can the entire repertoire be tapped?

One strategy that shows promise is to bypass the inefficient fusion step in hybridoma production and directly clone antibody cDNAs from the lymphocytes of immunized mice (Figures 23-6 and 23-7). Investigators inoculated a mouse with an antigen. They recovered spleen cells from the mouse and used PCR to amplify millions of cDNAs for antibody light and heavy chains. The light- and heavy-chain cDNAs were cloned separately into phage vectors and then recombined in vitro to generate a third, *combinatorial* library of phage carrying random pairs of light and heavy chains. The library was plated onto a bacterial lawn, and the resulting phage plaques, each containing a unique antibody, were screened with radioactively labeled antigen in a manner similar to that used for

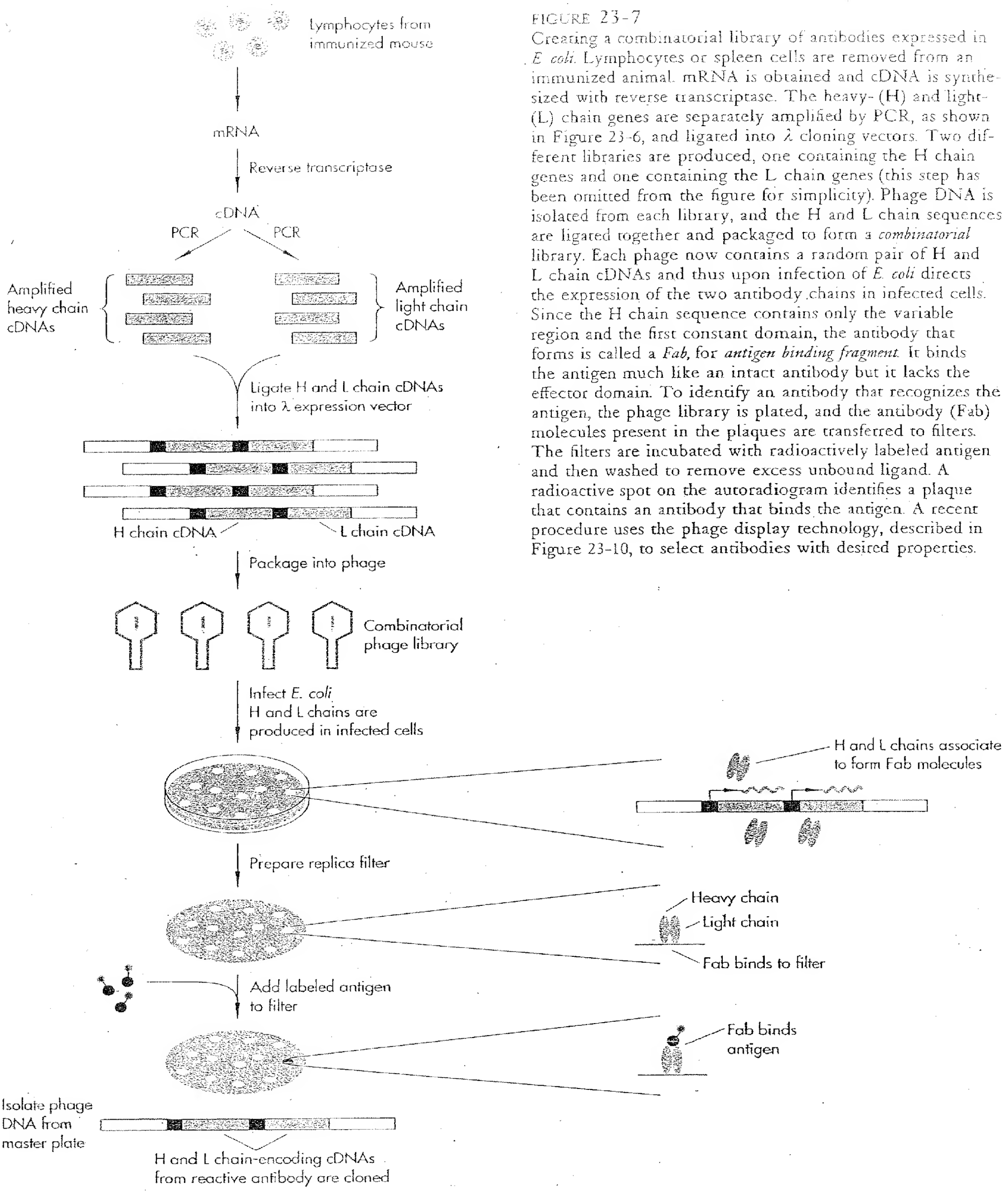


FIGURE 23-7
Creating a combinatorial library of antibodies expressed in *E. coli*. Lymphocytes or spleen cells are removed from an immunized animal. mRNA is obtained and cDNA is synthesized with reverse transcriptase. The heavy- (H) and light- (L) chain genes are separately amplified by PCR, as shown in Figure 23-6, and ligated into λ cloning vectors. Two different libraries are produced, one containing the H chain genes and one containing the L chain genes (this step has been omitted from the figure for simplicity). Phage DNA is isolated from each library, and the H and L chain sequences are ligated together and packaged to form a *combinatorial* library. Each phage now contains a random pair of H and L chain cDNAs and thus upon infection of *E. coli* directs the expression of the two antibody chains in infected cells. Since the H chain sequence contains only the variable region and the first constant domain, the antibody that forms is called a *Fab*, for *antigen binding fragment*. It binds the antigen much like an intact antibody but it lacks the effector domain. To identify an antibody that recognizes the antigen, the phage library is plated, and the antibody (Fab) molecules present in the plaques are transferred to filters. The filters are incubated with radioactively labeled antigen and then washed to remove excess unbound ligand. A radioactive spot on the autoradiogram identifies a plaque that contains an antibody that binds the antigen. A recent procedure uses the phage display technology, described in Figure 23-10, to select antibodies with desired properties.

cloning cDNAs from an expression library (Figure 7-10). Out of a million phage plaques screened, 200 clones were identified that produced an antibody binding the antigen. Thus, with this approach, investigators were able to sample a million possible antibodies—at least a thousand times more than they could screen by conventional monoclonal antibody technology. Since phages in a particular plaque encode the antibody expressed in the plaque, it is a trivial matter to clone the heavy- and light-chain cDNAs from the phage DNA. These cDNAs can be placed into bacterial or mammalian expression vectors for production of large quantities of the selected antibody.

A recent modification of this method uses filamentous phages such as M13 instead of λ phage and allows display of the antibodies on the phage surface. This offers the advantage of being able to screen thousands more phage (because the screening can be done in solution) and to select phage that express tight-binding antibodies. We will discuss this method later and in Figure 23-10.

"Humanized" Monoclonal Antibodies Retain Activity But Lose Immunogenicity

Although swift progress is being made in the identification of monoclonal antibodies with potential therapeutic value, their use is limited by a problem we have already discussed in this chapter. Monoclonal antibodies are usually mouse proteins, and they are not identical to human antibodies. Thus, antibodies injected into a patient will eventually be recognized as foreign proteins and will be cleared from the circulation.

As we learned in Chapter 16, both chains of the antibody molecule can be divided into variable and constant regions. The variable regions differ in sequence from one antibody to another, and this is the region of the protein that binds the antigen. The constant region is the same among all antibodies of the same type. The first method used to reduce the immunogenicity of a mouse monoclonal antibody was simply to construct *chimeric* genes that encoded proteins in which the variable regions from the mouse antibody were fused to the constant regions from a human antibody. The chimeric antibody (Figure 23-8) retained its binding specificity but more closely resembled a natural human antibody.

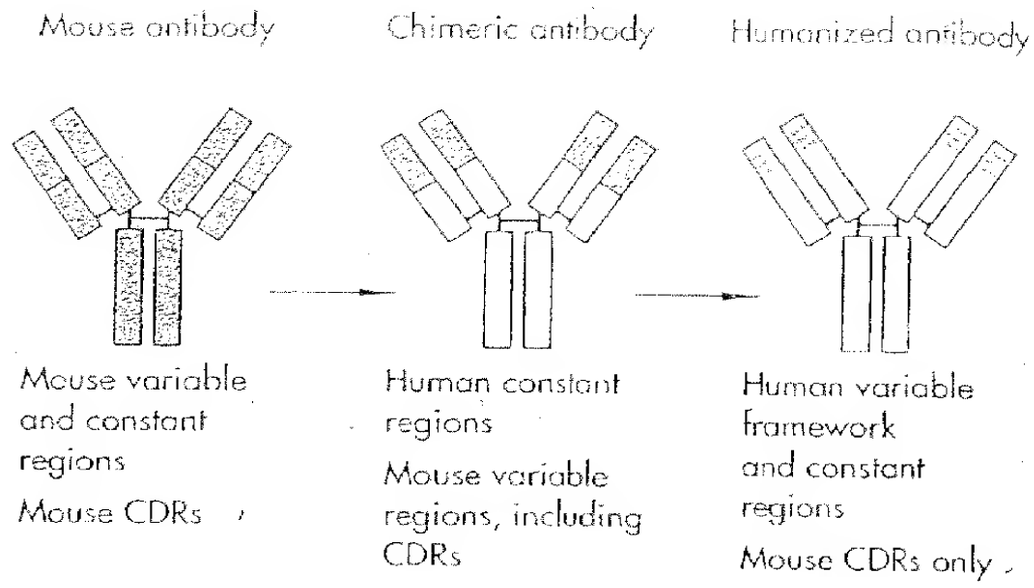


FIGURE 23-8
Antibody engineering. The basic structure of a mouse monoclonal antibody (MAb) resembles that of a human antibody. However, there are numerous differences between amino acid sequences of the antibodies from the two species. These sequence differences account for the immunogenicity of mouse MAbs in humans. A *chimeric* MAb is constructed by ligating the cDNA fragment encoding the mouse V_L and V_H domains to fragments encoding the C domains from a human antibody. Because the C domains do not contribute to antigen binding, the chimeric antibody will retain the same antigen specificity as the original mouse MAb but will be closer to human antibodies in sequence. Chimeric MAbs still contains some mouse sequences, however, and may still be immunogenic. A *humanized* MAb contains only those mouse amino acids necessary to recognize the antigen. This product is constructed by building into a human antibody the amino acids from the mouse complementarity determining regions or CDRs.

This antibody, however, was not fully *humanized*, because it retained amino acid sequences from the mouse protein. Thus, scientists have set out to engineer fully humanized monoclonal antibodies that will be indistinguishable from natural molecules. Extensive studies of the three-dimensional structures of antibody molecules tell us that only a few of the one hundred amino acids in the variable region of an antibody actually contact the antigen; these regions of contact are referred to as *complementarity determining regions* (CDRs). Three CDRs each comprise the antigen-binding sites on the light and heavy chains. The rest

of the variable region serves as a scaffold to anchor the CDRs in the correct positions. This breakdown of amino acids in the variable region into those serving recognition and those serving structural roles is also evident from simply comparing the sequences of many antibody molecules. Amino acid sequences in the CDRs are *hypervariable,* whereas the structural, or *framework,* amino acids differ little.

Thus, to make a fully humanized antibody, all that would be required in principle would be to use in vitro mutagenesis to transfer the CDR amino acid sequences from a mouse MAb to a natural human antibody (Figure 23-8). This method was used to humanize an antibody that recognizes an antigen on the surface of human lymphocytes. This humanized MAb is now in clinical trials as an immunosuppressant and for treatment of lymphoid tumors. Another potentially valuable MAb binds a growth factor receptor found in large numbers on the surface of many breast tumor cells. Laboratory experiments showed that this antibody could block the growth of these cells in culture and caused tumors seeded in mice to regress. Unfortunately, the first humanized versions of this antibody bound the receptor protein but failed to block the growth of breast carcinoma cells. Investigators suspected that the problem was with the framework amino acids, and they used computer modeling to design amino acid substitutions that would strengthen the antibody-antigen interaction. Several such variant antibodies were produced and tested; one bound the receptor 250 times more tightly than did the original antibody and successfully blocked tumor cell growth in culture. This antibody is now being produced in large quantities for clinical trials.

Protein Engineering Can Tailor Antibodies for Specific Applications

Humanizing monoclonal antibodies is an example of the emerging technology of *protein engineering,* that is, a process using recombinant DNA to modify the structure of natural proteins to improve or change their function. Antibodies are particularly attractive candidates for protein engineering, because their structure is understood in great detail and because their potential for use in medicine is enormous. Another way in which antibodies are being engineered is by changing their *effector domains,* the regions of the heavy chain that specify antibody function—for example, killing of cells marked by the antibody. In this way, the mode of action of a monoclonal antibody can be reprogrammed. One promising strategy is to replace the effector domain entirely with a sequence encoding a toxin. An antibody-toxin fusion protein would deliver the toxin specifically to cells bearing the target antigen. This product could be an exceptionally potent treatment for cancer and for viral diseases such as AIDS. Antibody engineering is also being used to construct *bispecific antibodies.* In these antibodies, each of the two arms recognizes a different antigen, thus allowing an antibody to bridge the two antigens. For example, a bispecific antibody could recognize a tumor cell protein with one arm and a protein on the surface of a killer T cell with the other, thereby bringing the killer cells directly to the tumor (Figure 23-9).

Protein Engineering Is Used to Improve a Detergent Enzyme

Subtilisin is a serine protease produced by bacteria. Due to its broad specificity for proteins that commonly soil clothing, this enzyme was developed for commercial use in laundry detergents. (It is subtilisin that is prominently advertised as the enzyme additive in modern detergents.) But the first detergents containing subtilisin suffered from a serious drawback: they could not be used with bleach, because bleach inactivates the enzyme. Biochemical analysis determined that loss of activity was due to the oxidation of a methionine at position 222. Once this happened, the modified enzyme lost 90 percent of its activity. Because they knew which amino acid was bleach sensitive, however, scientists decided to see whether a variant of subtilisin could be produced that was no longer sensitive to bleach.

To do this, site-directed mutants were constructed in the gene encoding subtilisin. The strategy was simply to substitute, one at a time, each of the non–wild-type amino acids at residue 222. The mutant genes

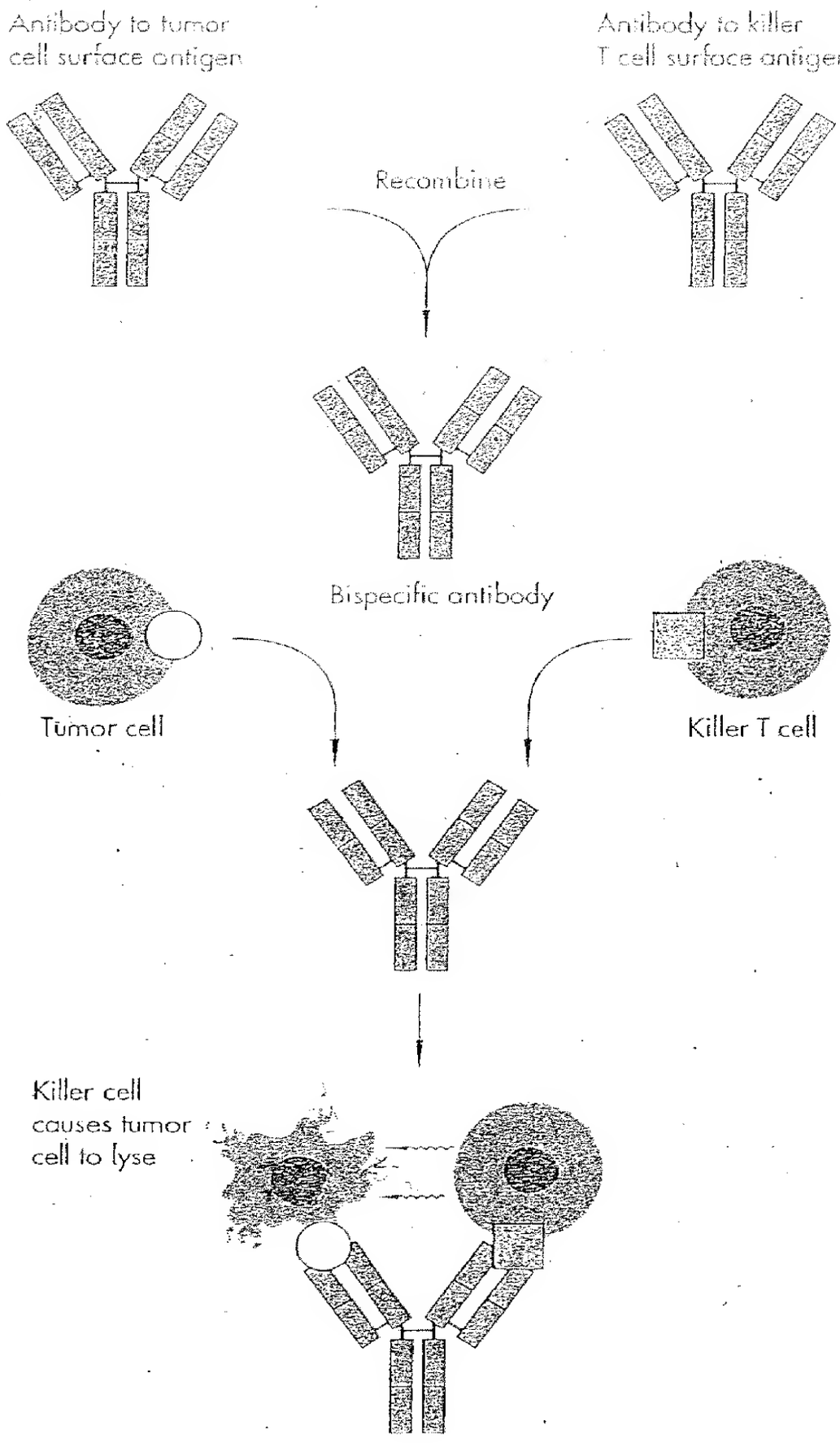


FIGURE 23-9
A bispecific antibody. By using recombinant DNA, the cDNAs for antibodies to two different antigens can be engineered to make an antibody in which each arm recognizes a different antigen. Thus it is possible to recombine antibodies to surface antigen on tumor cells and to a protein on cytotoxic T cells to make a bispecific antibody that brings the two cells together to facilitate killing of the tumor cells.

were cloned into expression vectors and the 19 different subtilisin derivatives were expressed. Biochemical analysis showed that the cysteine-222 enzyme was even more active than the wild-type protein, but it was also inactivated by bleach. The next most active variant was the alanine-substituted enzyme, which was 53 percent as active as wild-type subtilisin. This variant exhibited no detectable bleach sensitivity, so detergents containing this engineered subtilisin can now be used with bleach. This new variant of subtilisin is an example of a *second-generation* molecule, a molecule specifically engineered for a new desirable trait. Protein engineers are currently at work on a third-generation molecule that exhibits decreased temperature sensitivity so that it can be used in hot water.

This experiment points out the power of recombinant DNA as a tool for the engineering of natural products. Changing the properties of a protein was all but impossible prior to the development of recombinant DNA techniques. Now it is not only possible, but easy. It is a routine exercise for protein engineers to generate hundreds of variants of a natural protein for testing. These changes can be educated guesses based on detailed knowledge of the structure of a protein; alternatively, changes can easily be made on a purely random basis. And, as we will see in the next section, a combination of structural information with random mutagenesis and a powerful selection for improved protein function can have dramatic results.

Growth Hormone Variants with Improved Binding Are Selected by Phage Display

To engineer an improved subtilisin enzyme, researchers were aided by the knowledge that only one specific amino acid had to be changed. Thus, they could systematically vary that amino acid to find the one that worked the best. But more complex challenges face protein engineers. Is it possible, for example, to engineer antibodies with higher affinity for antigen; to design an inhibitor that tightly binds to and blocks a cell-surface protein or an enzyme inside a cell; to generate a growth factor or hormone with increased affinity for its receptor? Alterations of this sort require several amino acid changes, and with 20 possible amino acids at each position, the number of variants that

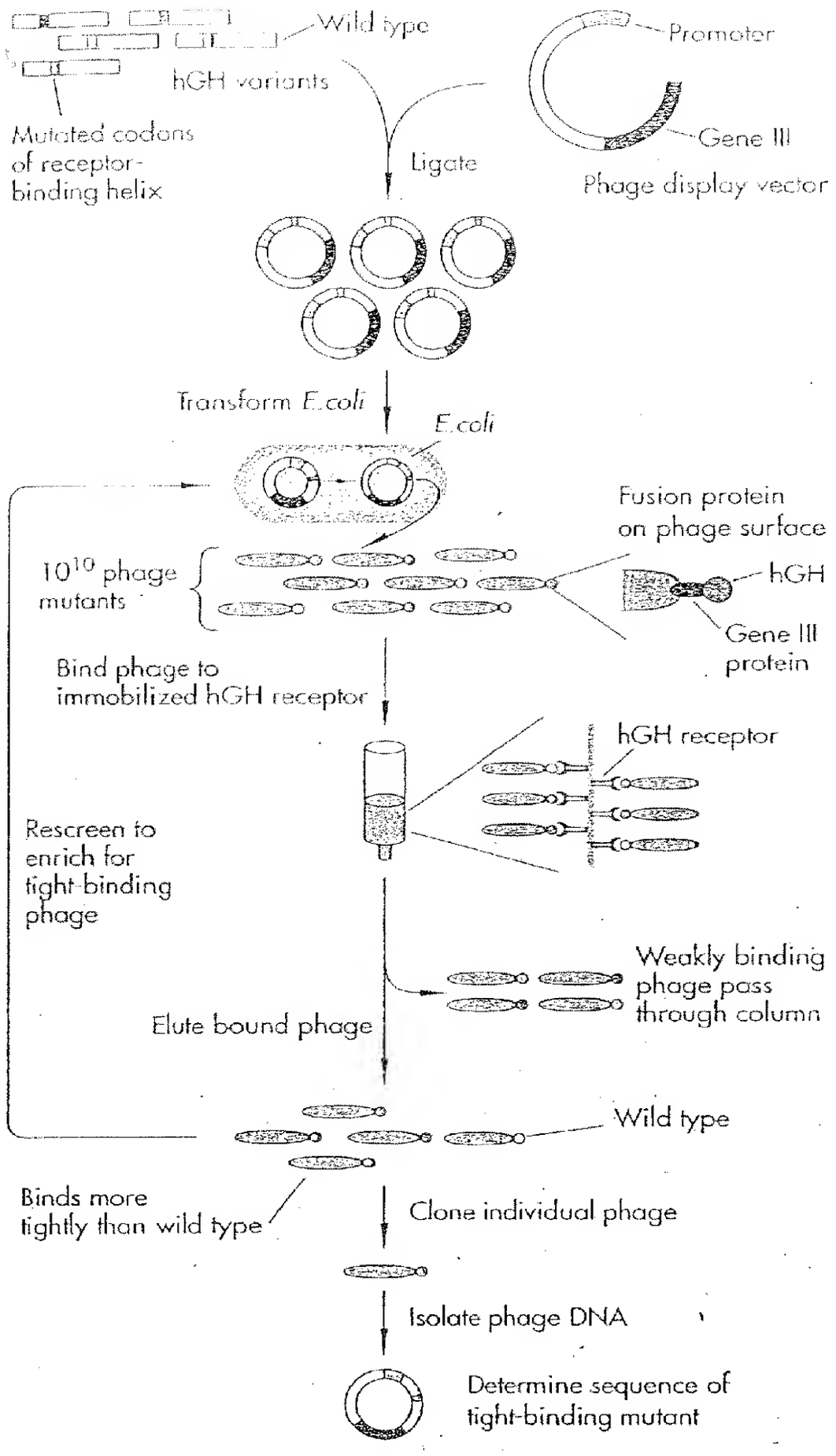


FIGURE 23-10
Expression of proteins and peptides on the surface of filamentous phage. A library of randomly mutated hGH cDNAs was ligated into an M13-based phagemid vector so that hGH was fused to the carboxy-terminal domain of the M13 gene III protein. The carboxy terminus of the gene III protein associates with the phage particle, and the amino terminus, containing the hGH variants, is displayed on the outer surface of the phage. The library of phagemids is introduced into *E. coli*, and ampicillin-resistant colonies are obtained. These *E. coli* are then infected with a helper phage that induces the production of phagemid particles. Only 1–10 percent of the phage particles contain an hGH–gene III fusion protein, and these contain only one hGH fusion molecule per phage. This ensures that the phage retain sufficient wild-type gene III protein in their coats to remain infectious. hGH-phage were passed through a column containing the hGH receptor covalently linked to plastic beads. Only the phage expressing hGH were retained. The nonbinding phage lacking hGH passed through the column. The bound phage were isolated, cultured in *E. coli*, and passed again over the column. Repeated rounds of selection resulted in the identification of hGH variants that bound the receptor with exceptionally high affinity.

need to be screened is enormous (for changes at just 3 amino acids, there are 8000 different combinations; for 10 amino acids, 10^{13} different proteins are possible). Clearly, these variants cannot be made and tested one at a time, and a method for direct selection of improved proteins is needed.

Researchers have used a new approach to select variants of human growth hormone with increased affinity for growth hormone receptor (Figure 23-10). From structural studies and extensive mutagenesis of hGH, they knew what portions of the amino acid sequence were important for receptor binding. They synthesized degenerate oligonucleotides that encoded all possible amino acids at these positions and ligated the pool of oligonucleotides in place of the natural hGH sequence. The resulting pool of variant hGH cDNAs was fused to the reading frame of gene III in the filamentous phage M13. Gene III encodes a minor phage coat protein expressed on the surface of the phage, and incorporation of the hGH cDNA into this gene results in the display of the hGH variants on the surface of the phage, one variant per phage. This technique is known as *phage display*.

Now it was a simple matter to pass this library of more than 10^{11} different phage over a column containing the hGH receptor. Phage displaying weakly binding hGH variants were washed off the column, and phage displaying tightly binding variants were recovered with a more stringent wash. This population of tight-binding phage was amplified by infection of *E. coli* and passed over the column a second time. The selection was repeated for a total of six rounds, each round enriching for the phage displaying hGH variants

with highest affinity for the receptor bound to the column. At this point, individual phage were cloned, the affinities of their hGH variants were measured directly, and the sequences of the hGH cDNAs were examined. Among these variants was one that bound its receptor about 10 times more tightly than natural hGH did. When selected amino acids from another region of hGH that had been randomized were introduced into this variant, the resulting hGH molecule bound to the hGH receptor over 50 times more tightly than the wild-type hGH did. This process is being repeated in the hope to obtain even more tightly binding variants.

The ability afforded by techniques such as phage display to correlate protein structure and function in a systematic way makes possible new methods of finding novel drugs. If researchers have a good idea what combination of amino acids gives the best fit to the binding site on a receptor, the next step in *rational drug design* would be to design, or even select, a small peptide that binds as well as the larger protein. And then, using computer modeling to display the molecular contacts between ligand and receptor, researchers can attempt to design and synthesize small nonprotein molecules that make the same contacts. The end-product would be a small organic molecule that could be produced more cheaply than a recombinant protein, yet would retain the full biological activity of the protein hormone. And, more important, such molecules could be administered orally, thus eliminating the major disadvantage of most recombinant protein therapeutics—that they must be delivered directly into the bloodstream by injection. This type of rational drug design contrasts sharply with the conventional approach to drug discovery now in use in the pharmaceutical industry, in which an inventory of completely unrelated compounds is tested at random until an active compound is found.

New Technologies Promise New Approaches to Drug Design

The biotechnology industry is in its infancy, and its successes to date follow directly from developments in molecular biology that are already nearly two decades old. The recombinant drugs currently in clinical use arise from what is by now conventional technology—gene cloning, expression, and mutagenesis to improve protein function. These methods will continue to turn out new drugs such as erythropoietins to treat anemia caused by kidney disease, DNase to treat cystic fibrosis, or colony-stimulating factors (CSFs) to increase white blood cell production during chemotherapy.

But the true promise of biotechnology is in novel technologies that are only now being developed. We have mentioned efforts to design catalytic antibodies that can accelerate chemical reactions in both medical and industrial applications. This is but one example of a whole new approach to protein engineering in which novel activities can be placed on unrelated protein scaffolds, using random mutagenesis coupled with selection methods like phage display. Similar goals may be achieved by the engineering of *ribozymes*, RNA molecules with catalytic activity, and the use of the polymerase chain reaction to select nucleic acid molecules that bind tightly to targets of medical importance. Another strategy that may see widespread application is treatment with antisense DNA and RNA to inhibit the expression of oncogenes in tumors or of viral genes in infected patients. And a variety of new technologies based on viral vectors promise new approaches for vaccines and gene therapy.

Many of these techniques now work in the test tube, and the principal challenge facing biotechnology companies is to turn these laboratory techniques into commercially viable processes.

Reading List

General

Hall, S. S. *Invisible Frontiers: The Race to Synthesize a Human Gene.* Atlantic Monthly Press, New York, 1987.

Hood, L. "Biotechnology and medicine of the future." *J. Am. Med. Assoc.*, 259: 1837–1844 (1988).

Goeddel, D. V. (ed.) Systems for Heterologous Gene Expression. *Meth. Enzymol.*, Vol. 185, Academic Press, New York, 1990.

Original Research Papers

EXPRESSION OF HUMAN PROTEINS IN E. COLI

Itakura, K., T. Hirose, R. Crea, A. Riggs, H. L. Heyneker, F. Bolivar, and H. Boyer. "Expression in *E. coli* of a chemically synthesized gene for the hormone somatostatin." *Science*, 198: 1056–1063 (1977).

Goeddel, D. V., H. L. Heyneker, T. Hozumi, R. Arentzen, K. Itakura, D. G. Yansura, M. J. Ross, G. Miozzari, R. Crea, and P. Seeburg. "Direct expression in *Escherichia coli* of a DNA sequence coding for human growth hormone." *Nature*, 281: 544–548 (1979).

Goeddel, D. V., D. G. Kleid, F. Bolivar, H. L. Heyneker, D. G. Yansura, R. Crea, T. Hirose, A. Kraszewski, K. Itakura, and A. D. Riggs. "Expression of chemically synthesized genes for human insulin." *Proc. Natl. Acad. Sci. USA*, 76: 106–110 (1979).

EXPRESSION IN YEAST

Valenzuela, P., A. Medina, W. J. Rutter, G. Ammerer, and B. D. Hall. "Synthesis and assembly of hepatitis B virus surface antigen particles in yeast." *Nature*, 298: 347–350 (1982).

Hirzeman, R. A., D. W. Leung, L. J. Perry, W. J. Kohr, H. L. Levine, and D. V. Goeddel. "Secretion of human interferons by yeast." *Science*, 219: 620–625 (1983).

Sabin, E. A., C. T. Lee-Ng, J. R. Shuster, and P. J. Barr. "High-level expression and in vivo processing of chimeric ubiquitin fusion proteins in *Saccharomyces cerevisiae*." *Bio/Tech.*, 7: 705–709 (1989).

EXPRESSION IN INSECT CELLS

Luckow, V. A., and M. D. Summers. "Trends in the development of baculovirus expression vectors." *Bio/Tech.*, 6: 47–55 (1988). [Review]

Medin, J. A., L. Hunt, K. Gathy, R. K. Evans, and M. S. Coleman. "Efficient, low-cost protein factories: expression of human adenosine deaminase in baculovirus-infected insect larvae." *Proc. Natl. Acad. Sci. USA*, 87: 2760–2764 (1990).

EXPRESSION OF PROTEINS IN MAMMALIAN CELLS

Gorman, C. M. "Mammalian cell expression." *Curr. Opin. Biotech.*, 1: 36–43 (1990). [Review]

Pennica, D., W. E. Holmes, W. J. Kohr, R. N. Harkins, G. A. Vehar, C. A. Ward, W. F. Bennett, E. Yelverton, P. H. Seeburg, H. L. Heyneker, D. V. Goeddel, and D. Collen. "Cloning and expression of human tissue-type plasminogen activator cDNA in *E. coli*." *Nature*, 301: 214–221 (1983).

Paborsky, L. R., B. M. Fendly, K. L. Fisher, R. M. Lawn, B. J. Marks, G. McCray, K. M. Tate, G. A. Vehar, and C. M. Gorman. "Mammalian cell transient expression of tissue factor for the production of antigen." *Prot Eng*, 3: 547–553 (1990).

VACCINES

Brown, F. "From Jenner to genes—the new vaccines." *Lancet*, 335: 587–590 (1990). [Review]

Bolognesi, D. P. "Approaches to HIV vaccine design." *Trends Biotech.* 8: 40–45 (1990). [Review]

Berman, P. W., T. J. Gregory, L. Riddle, G. R. Nakamura, M. A. Champe, J. P. Porter, F. M. Wurm, R. D. Hershberg, E. K. Cobb, and J. W. Eichberg. "Protection of chimpanzees from infection by HIV-1 after vaccination with recombinant glycoprotein gp120 but not gp160." *Nature*, 345: 622–625 (1990).

PROTEIN ENGINEERING RECOMBINANT PRODUCTS

Estell, D. A., T. P. Graycar, and J. A. Wells, "Engineering an enzyme by site-directed mutagenesis to be resistant to chemical oxidation." *J. Biol. Chem.* 260: 6518–6521 (1985).

DesJarlais, R. L., G. L. Seibel, I. D. Kuntz, P. S. Furth, J. C. Alvarez, P. R. Ortiz de Montellano, D. L. DeCamp, L. M. Babe, and C. S. Craik. "Structure-based design of nonpeptide inhibitors specific for the human immunodeficiency virus 1 protease." *Proc. Natl. Acad. Sci. USA*, 87: 6644–6648 (1990).

Abrahmsen, L., J. Tom, J. Burnier, K. A. Butcher, A. Kossiakoff, and J. A. Wells. "Engineering subtilisin and its substrates for efficient ligation of peptide bonds in aqueous solution." *Biochemistry*, 30: 4151–4159 (1991).

Bennett, W. F., N. F. Paoni, B. A. Keyt, D. Botstein, A. J. S. Jones, L. Presta, F. M. Wurm, and M. J. Zoller. "High resolution analysis of functional determinants on human tissue-type plasminogen activator." *J. Biol. Chem.*, 266: 5191–5201 (1991).

Cunningham, B. C. and J. A. Wells. "Rational design of receptor-specific variants of human growth hormone." *Proc. Natl. Acad. Sci. USA*, 88: 3407–3411 (1991).

MONOCLONAL ANTIBODIES

Milstein, C. "Monoclonal antibodies." *Sci. Am.*, 243: 66–74 (1980). [Review]

CLONING AND EXPRESSION OF ANTIBODIES

Pluckthun, A. "Antibodies from *Escherichia coli*." *Nature*, 347: 497–498 (1990). [Review]

Huse, W. D., L. Sastry, S. A. Iverson, A. S. Kang, M. Alting-Mies, D. R. Burton, S. J. Benkovic, and R. A. Lerner. "Generation of a large combinatorial library of the immunoglobulin repertoire in phage lambda." *Science*, 246: 1275–1289 (1989).

Chaudhary, V. K., J. K. Batra, M. G. Gallo, M. C. Willingham, D. J. FitzGerald, and I. Pastan. "A rapid method of cloning functional variable-region antibody genes in *Escherichia coli* as single-chain immunotoxins." *Proc. Natl. Acad. Sci. USA*, 87: 1066–1070 (1990).

Mullinax, R. L. et al. "Identification of human antibody fragment clones specific for tetanus toxoid in a bacteriophage λ immunoexpression library." *Proc. Natl. Acad. Sci. USA*, 87: 8095–8099 (1990).

Berg, J., E. Lotscher, K. S. Steimer, D. J. Capon, J. Baenziger, H. M. Jack, and M. Wabl. "Bispecific antibodies that mediate killing of cells infected with human immunodeficiency virus of any strain." *Proc. Natl. Acad. Sci. USA*, 88: 4723–4727 (1991).

Wood, C.R., A. J. Dorner, G. E. Morris, E. M. Alderman, D. Wilson, R. M. J. O'Hara, and R. J. Kaufman. "High level synthesis of immunoglobulins in Chinese hamster ovary cells." *J. Immunol.*, 145: 3011–3016 (1990).

CATALYTIC ANTIBODIES

Jacobs, J. W. "New perspective on catalytic antibodies." *Bio/Tech.*, 9: 258–262 (1991). [Review]

Benkovic, S. J., J. A. Adams, C. L. Borders, K. D. Janda, and R. A. Lerner. "The enzymic nature of antibody catalysis: development of multistep kinetic processing." *Science*, 250: 1135–1139 (1990).

Bowdish, K., Y. Tang, J. B. Hicks, and D. Hilvert. "Yeast expression of a catalytic antibody with chorismate mutase activity." *J. Biol. Chem.*, 266: 11901–11908 (1991).

HUMANIZING ANTIBODIES

Winter, G., and C. Milstein. "Man-made antibodies." *Nature*, 349: 293–299 (1991). [Review]

Jones, P. T., P. H. Dear, J. Foote, M. S. Neuberger, and G. Winter. "Replacing the complementarity-determining regions in a human antibody with those from a mouse." *Nature*, 321: 522–525 (1986).

Riechmann, L., M. Clark, H. Waldmann, and G. Winter. "Reshaping human antibodies for therapy." *Nature*, 332: 323–327 (1988).

Carter, P., L. Presta, C. M. Gorman, J. B. Ridgway, D. Henner, W. L. T. Wong, A. M. Rowland, C. Kotts, M. E. Carver, and M. H. Sheppard. "Humanization of an anti-p185^{HER2} antibody for human cancer therapy." *Proc. Natl. Acad. Sci. USA*, 89: In press (1992).

PHAGE DISPLAY

Smith, G. P. "Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface." *Science*, 228: 1315–1317 (1985).

Cwirla, S. E., E. A. Peters, R. W. Barrett, and W. J. Dower. "Peptides on phage: a vast library of peptides for identifying ligands." *Proc. Natl. Acad. Sci. USA*, 87: 6378–6382 (1990).

Clackson, T., H. R. Hoogenboom, A. D. Griffiths, and G. Winter. "Making antibody fragments using phage display libraries." *Nature*, 352: 624–628 (1991).

Lowman, H. B., S. Bass, N. Simpson, and J. A. Wells. "Selecting high affinity binding proteins by monovalent phage display." *Biochemistry*, 30, 10832–10838 (1991).

NEW RECOMBINANT DRUGS

Shak, S., D. J. Capon, R. Hellmiss, S. A. Marsters, and C. L. Baker. "Recombinant human DNase I reduces the viscosity of cystic fibrosis sputum." *Proc. Natl. Acad. Sci. USA*, 87: 9188–9192 (1990).

Takaue, Y., et al. "Effects of recombinant human G-CSF, GM-CSF, IL-3, and IL-1 alpha on the growth of purified human peripheral blood progenitors." *Blood*, 76: 330–335 (1990).

Shepard, H.M., G. D. Lewis, J. C. Sarup, B. M. Fendly, D. Maneval, J. Mordenti, I. Figari, C. E. Kotts, M. A. Palladino, A. Ullrich, and D. slamon. "Monoclonal antibody therapy of human cancer: taking the HER2 protooncogene to the clinic." *J. Clin. Immunol.*, 11: 117–127 (1991).

Watson, S. R., C. Fennie, and L. A. Lasky. "Neutrophil influx into an inflammatory site inhibited by a soluble homing receptor-IgG chimaera." *Nature*, 349: 164–167 (1991).

FUTURE TRENDS

Sarver, N., E. M. Cantin, P. S. Chang, J. A. Zaia, P. A. Ladne, D. A. Stephens, and J. J. Rossi. "Ribozymes as potential anti-HIV-1 therapeutic agents." *Science*, 247: 1222–1225 (1990).

Tuerk, C., and L. Gold. "Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase." *Science*, 249: 505–510 (1990).

Han, L., J. S. Yun, and T. E. Wagner. "Inhibition of Moloney murine leukemia virus-induced leukemia in transgenic mice expressing antisense RNA complementary to the retroviral packaging sequences." *Proc. Natl. Acad. Sci. USA*, 88: 4313–4317 (1991).

Miller, P. S. "Oligonucleoside methylphosphonates as antisense reagents." *Bio/Tech.*, 9: 358–362 (1991).

Schultz, J. S. "Biosensors." *Sci. Am.*, 265: 64–69 (1991).